DTIC FILE COPY

2

Report AFOSR-85-0370

AFOSR-TR- 89-0769

AD-A208 695

Perception of Motion in Statistically-Defined Displays

Robert Sekuler
Cresap Laboratory
Northwestern University
Evanston IL 60201

15 April 1989

Final Scientific Report for Period
1 October 1985 to 30 September 1988

Prepared for
Air Force Office of Scientific Research
Bolling Air Force Base, DC 20332

89 6 06 071

Report AFOSR-85-0370

Perception of Motion in Statistically-Defined Displays

Robert Sekuler
Cresap Laboratory
Northwestern University
Evanston IL 60201

15 April 1989

Final Scientific Report for Period
1 October 1985 to 30 September 1988

Prepared for
Air Force Office of Scientific Research
Bolling Air Force Base, DC 20332

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Field	Entry
1a. REPORT SECURITY CLASSIFICATION	Unclassified
1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
3. DISTRIBUTION/AVAILABILITY OF REPORT	Unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	
5. MONITORING ORGANIZATION REPORT NUMBER(S)	AFOSR-TR- 89-0769
6a. NAME OF PERFORMING ORGANIZATION	Northwestern University
6b. OFFICE SYMBOL (If applicable)	
7a. NAME OF MONITORING ORGANIZATION	Air Force Office of Scientific Research NL
6c. ADDRESS (City, State and ZIP Code)	Psychology Department, 102 Swift Hall, Evanston, IL 60208
7b. ADDRESS (City, State and ZIP Code)	Bolling Air Force Base, Washington, D.C. 20332-6448 Bldg 410
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	AFOSR
8b. OFFICE SYMBOL (If applicable)	NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	AFOSR 85-0370
8c. ADDRESS (City, State and ZIP Code)	Bldg 410, Bolling AFB DC 20332-6448
10. SOURCE OF FUNDING NOS.	PROGRAM ELEMENT NO.: 61102F; PROJECT NO.: 2313; TASK NO.: A5; WORK UNIT NO:
11. TITLE (Include Security Classification)	Perception of Motion in Statistically-Defined Displays
12. PERSONAL AUTHOR(S)	Sekuler, Robert
13a. TYPE OF REPORT	Final Technical
13b. TIME COVERED	FROM 85/10/1 TO 88/9/30
14. DATE OF REPORT (Yr., Mo., Day)	89/4/15
15. PAGE COUNT	272
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES (FIELD / GROUP / SUB. GR.)	
18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	Vision, perception, motion, mathematical model, computational model, statistical displays, cinematograms

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Overview. During the three-year reporting period covered by this Final Scientific Report, we carried out a series of coordinated, major projects. Like the predecessor project (AFOSR 80-0246), work during this period used statistically-complex displays to probe higher-order of human motion perception. The project's overall aim was to establish the characteristics of human visual mechanisms that extract information from these and other displays. A fundamental contribution was the systematic refinement and extension of a model in which motion information is extracted and processed, via non-linear interactions, by directionally-selective visual mechanisms.

Continued on reverse....

Field	Entry
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐
21. ABSTRACT SECURITY CLASSIFICATION	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	Dr. John F. Tangney
22b. TELEPHONE NUMBER (Include Area Code)	(202) 767-5021
22c. OFFICE SYMBOL	NL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

Block 19, Continued

Project One. Prior work in our laboratory showed that a percept of global coherent motion can be produced from the combination of many different, localized motion vectors. Now, using random-dot cinematograms, we established that hysteresis is strongly associated with such percepts. D.W. Williams, G. Phillips and R. Sekuler showed that the characteristics of the hysteresis are relatively robust with respect to changes in dot density, display area and location. Changing the display's directional content, however, did alter the hysteresis profile in a manner that is consistent with a model incorporating cooperative interactions among direction-selective motion mechanisms. These results lend significant support to a view of motion processing in which cooperative interactions play a prominent role.

Project Two. A second major project during this period followed-up our previous finding that practice seemed to produce direction-selective improvement in observers' ability to discriminate between highly similar directions of motion. Kosnik, Fikre and Sekuler clarified the basis for this improvement by recording an observers eye movements while they tried to discriminate between slightly different directions of target motion. We found that observers did not need to track the moving target in order to learn the discrimination. These results suggest that practice's influence on the discrimination of motion's direction is perceptual rather than sensori-motor in character.

Project Three. S. Watamaniuk, R. Sekuler and D.W. Williams created random-dot cinematograms in which each dot's successive movements were independently drawn from a Gaussian distribution of directions of some characteristic bandwidth. As established earlier, such displays, comprising many different, spatially intermingled local motion vectors, can produce a percept of global coherent motion in a single direction. Using pairs of cinematograms, direction discrimination of global motion was measured under various conditions of direction distribution bandwidth, exposure duration, and constancy of each dot's path. A line-element model gave an excellent account of the results: i) over a considerable range, discrimination was unaffected by the cinematogram's direction distribution bandwidth; ii) only for the briefest presentations did changes in duration have an effect; iii) so long as the overall directional content of the cinematogram remained unchanged, the constancy or randomness of individual dots' paths did not affect discrimination. Finally, the line-element model continued to give a good account of the results when we made additional measurements with uniform rather than Gaussian distributions of directions.

Project Four. This project extended previous work on the

AFOSR-TR- 89-0769

perception of motion direction and speed to an important related case, perception of *change* in velocity. E. Dzhafarov and R. Sekuler set out to identify the information that controlled speeded response to motion onset or change in motion. Observers were required to react to the change in movement of a random-dot field whose velocity switched abruptly from V_0 to V_1. Changes in velocity were created by either shifting the speed, with direction constant, or by reversing direction, with speed constant. Mean reaction times and their standard deviations were decreasing functions of the difference $|V_1-V_0|$ and increasing function of the initial speed, $|V_0|$. The results are quantitatively accounted for by a modification of the Local Dispersion model that Dzhafarov and J. Allik proposed for motion detectability. In our modification, detection of change of velocity from V_0 to V_1 is treated as structurally equivalent to the detection of onset of a motion whose velocity is $|V_1-V_0|$. We have found that the Local Dispersion model can be realized by the mass activation of network of simple, bilocal correlators, like those proposed by J. Koenderink.

Project Five. M. Nawrot and R. Sekuler used random-dot cinematograms to examine how motion within one region of space influences the motion seen in another, neighboring region. The cinematograms were spatially heterogeneous, comprising alternating strips within which dots i)tended to move in one direction, or ii)moved about randomly (dynamic noise). When the alternating strips were narrow, motion in one direction induced a similar direction of illusory motion in the adjoining dynamic noise (assimilation); when alternating strips were wide, motion tended to induce an illusory opposed motion in the dynamic noise (contrast). Since it exhibits hysteresis, this illusory motion probably results from a network of spatially distributed, cooperative processes. The shift from assimilation to contrast, as the cinematogram's strips increase in size, suggests that facilatory and inhibitory influences of the network extend over different distances. To account for these results, required only a small addition to the model proposed earlier in this reporting period by Williams, Sekuler, and Phillips.

Project Six. D.W. Williams and G. Phillips extended our earlier work on random-dot cinematograms to the domain of three-dimensional structure from motion. It's been long known that the human visual system can recover the correct three-dimensional structure of moving objects solely from the relative changes in the two-dimensional retinal projection. The basis for this ability is unclear since infinitely many combinations of three- dimensional structure and motion can project to the same two-dimensional image. Using a stochastic random-dot cinematogram, Williams and Phillips demonstrated

Block 19, Continued

that the recovery of structure from motion does not depend upon the details of the spatio-temporal relations among elements of the image, but rather upon the overall directional content of the motion in the image. Further, the three-dimensional percept obtained with the random-dot stimulus exhibits hysteresis behavior. Changing the directional content of the stimulus altered the hysteresis profile in manner consistent with the Williams et al. model (developed earlier in this period) incorporating cooperative interactions among direction-selective mechanisms. In addition, the results strongly challenge widely-held views of the recovery of structure from motion, including models that depend upon constraints such as rigidity or incremental rigidity.

Project Seven. As an ancillary to the experimental and theoretical work of the Projects One through Six, R. Sekuler organized a session on motion perception at the Badenweiler (West Germany) Conference on the physiological underpinnings of perception. He subsequently had sole responsibility for preparing a written version of that session. For the sake of completeness, that written version, which will appear as a chapter in a book to be published this year, is included in the report.

<u>Project One:</u>

Cooperative phenomena in the perception of motion direction

Douglas Williams, Gregory Phillips, and Robert Sekuler

INTRODUCTION

A collection of localized motion vectors can produce a percept of global coherent motion along a single direction, even though the directional range of the individual motion vectors is quite broad.[1] Chang and Julesz[2] have suggested that this coherent motion percept may reflect an underlying cooperative process. In general, a cooperative system consists of local elements that interact with each other, thereby generating global behavior that would not occur were the elements isolated from one another. One signature of a cooperative system, hysteresis, is a form of memory in which a system, having reached a stable state, shows resistance to further change. A consequence of such behavior is that the system's response depends upon the history of stimulation.

The first evidence of a cooperative phenomenon in the visual system was that of binocular neural hysteresis.[3] These authors found that, while it was necessary to bring a pair of random dot stereograms to within 6' visual angle of each other for stereoscopic fusion to occur, it was possible to pull the pair apart by as much as 2^{o} before fusion was lost. Once fusion was lost, the stereo pair had to be returned to a disparity of 6' for fusion to be reestablished. The amount of disparity required to fuse or split apart the two stereograms thus depended upon the initial perceptual condition and the direction of the disparity change.

In this paper, we seek to strengthen a cooperative interpretation of motion perception by looking for evidence of hysteresis in the perception of motion direction using random-dot cinematograms. In our cinematograms, each dot takes an independent, two-dimensional random walk of constant step size. Specifically, each dot's direction of displacement from one frame to the next is chosen randomly from a uniform distribution of directions. The percept that results depends upon the range of this uniform distribution.[1] For a range of 360°, only the local random motion of the individual dots is evident. For a range of 180° or less, however, the percept is that of global coherent motion along the direction of the mean of the distribution, although the individual perturbations of the dots are still evident.

If this percept of global coherent motion is a result of cooperative processing, one might then expect the percept to exhibit hysteresis behavior. That is, by gradually changing the directional content of the stimulus between the two extremes of a uniform distribution with range 180° or less and a uniform distribution with range 360°, one can measure the transition points marking the change from global coherent motion to local random motion and vice versa. The results would be indicative of hysteresis if the directional content of the stimulus for which these transitions occur depended upon whether the perceptual change was from local to global motion or from global to local motion.

Our experimental results confirm the existence of hysteresis for the global coherent motion percept. Furthermore, we have found it possible to account for this hysteresis by cooperative, nonlinear excitatory and inhibitory interactions among direction-selective mechanisms for motion. Other models incorporating such cooperative interactions have been successful in describing binocular stereopsis.[4-9]

METHODS

Our stimuli, dynamic random-dot cinematograms, were generated by a PDP 11/34 computer that passed values through a digital-to-analog converter for display on a Hewlett Packard 1321A X-Y display (P31 phosphor). A "wrap-around" scheme caused dots that were displaced beyond the boundary of the display to reappear at the opposite side of the display. A cardboard mask restricted the visible pattern to a circular region with a diameter of 16°. In the absence of a fixation point, observers were instructed to maintain their fixation at the center of the screen; viewing was monocular, with the other eye occluded by an opaque eye patch.

A detailed discussion of the spatial and temporal parameters of the display can be found in Williams and Sekuler.[1] However, to summarize briefly: the display was composed of 512 dots, each measuring 0.1° in diameter with a spatial density of approximately 1.6 dots/deg^2. From one display frame to the next, each dot was displaced by 0.9°. The frame duration -- the time required to present all the dots once -- was 9.0 msecs, with an interframe interval of 95.0 msecs.

The display itself provided the only luminance in the room and observers adapted to the light level of a blank screen for five minutes before starting an experimental session. At the beginning of each session, the threshold luminance for a field of stationary dots was established using a von Bekesy tracking procedure.[10] Thereafter, the cinematograms were presented at twice this threshold luminance.

Each dot in the display took a two-dimensional random walk of constant step size, drawing its direction of movement randomly from a uniform distribution of directions. As a result, a dot's direction of movement was independent not only of the displacements of the other dots in the field but also of its own prior displacements. Such a stimulus can generate a percept of global motion, depending upon the range of the underlying directional distribution. When the range of the distribution extends over a full 360°, the percept is that of only localized random motion, whereas if the distribution extends over a much smaller range of 180° or less, the percept is that of global coherent motion along the direction of the mean - upwards in our case.[1] Because of the percept associated with each, the 360° distribution is referred to as the "noise" stimulus, and the 180°, or smaller, distribution as the "signal".

During an experimental session two modes of presentation were randomly intermixed. In one mode, all dots initially chose their directions of movement from the signal distribution, corresponding to an initial percept of global coherent motion along the upward direction. After a random period of time lasting up to 12 seconds, the proportion of dots choosing their direction from the signal distribution slowly decreased by two dots per frame; in other words, the proportion of dots choosing from the noise distribution increased by two dots per frame. The observers were asked to report, by means of a response switch, when the field of dots first appeared to exhibit only local random motion, having previously exhibited global upward motion. Following this response, which was recorded by the computer, the proportion of signal continued to decrease for a random time up to 6 seconds. At this point, the proportion of signal now increased by two dots per frame until the observer reported that the percept of global coherent motion had been restored. This second transition was recorded and the trial was terminated. The two random intervals that were incorporated into each trial ensured that the observer could not use temporal cues in deciding when a transition occurred. Also, the rate of change in the signal/noise proportion (i.e., 2 dots per frame) was chosen so that the stimulus duration was not too long nor the response resolution too coarse.

In order to make the stimulus more stochastic, every dot on each frame was permitted to choose its direction of motion from either the signal or noise distribution, irrespective of which distribution it had chosen from on previous frames. In particular, a dot had a probability equal to the proportion of signal dots of choosing its direction of motion from the signal distribution and a probability equal to the proportion of noise of choosing from the noise distribution.

In the other mode of presentation, the trial structure was the same except that the initial stimulus condition was that of all noise, with the first and second perceptual transitions going from local random to global coherent motion and back again.

We have chosen to parameterize a transition by the proportion of dots choosing their direction of motion from the signal distribution at the transition. In a single experimental session, 20 measurements of the signal proportion were made at each of the two perceptual transitions. A complete data set typically comprised the results obtained over 5 sessions. Data from two naive observers are reported.

RESULTS

The results for the two observers are shown in Figure 1 for the case in which the range of the signal distribution was 90° (*i.e.*, $+45^{\circ}$ to -45° about the vertical). The two possible motion percepts, "global upward flow" and "local random motion", are shown versus the proportion of dots choosing their direction of motion from the signal distribution. In each panel of the figure, the solid circle and associated error bar indicate the mean and standard deviation of the measurements of the signal proportion for the transition from local random motion to global upward flow. The open circle represents the data for the opposite transition from global upward flow to local random motion. We shall refer to the perceptual transition from local random motion to global upward flow as the 'global' transition and the transition in the opposite direction as the 'local' transition. For each observer these respective transition points are significantly different at the 5% level.

The arrows and solid lines schematically represent changes in the perceptual state as the proportion of signal dots is either increased or decreased. As indicated by the lower path in each panel, the percept of only local random motion is unchanged as the proportion of signal is increased, until this proportion reaches the value shown at G, at which point a 'global' transition occurs. The global percept persists for all larger values of the signal proportion. Conversely, the upper path indicates that in order to lose the global percept it is necessary to decrease the proportion of signal to the value at L, below which the local percept then prevails. The criterion for the existence of hysteresis is that the proportion of signal at point L must be less than that at point G. The results for both observers obviously satisfy this criterion. It should be noted that the hysteresis profile is shown as square-cornered for schematic purposes. Observers did comment, however, upon the abruptness of the perceptual transitions.

At this point, we sought to firmly establish the role of stimulus history in the observed hysteresis. To do so, it was necessary to rule out some other potential explanations. It is unlikely that the hysteresis simply reflects a response delay due to reaction time, since the mean time between the global and local transitions is 8.6 sec for observer JF and 8.9 sec for TKD-- more than an order of magnitude greater than typical reaction times. Another potential explanation, the motion aftereffect, would actually tend to diminish the width of the hysteresis profile, since it would likely hasten, and not retard, the perceptual transition from upward flow to local,random motion.

As a consequence, it is possible that the hysteresis may be even more robust than we have observed. Lastly, our results might be complicated by eye movements. Because of the stochastic nature of the stimulus, it would be difficult to track individual dots; however, eye movements could be entrained to the upward flow. To examine this possibility we repeated the experiments with a fixation dot in the center of the screen. Eye movement recordings obtained by Kosnik et al.[11] indicate that the directions of eye movements are not correlated with the direction of movement in random dot stimuli if a stationary fixation dot is provided. Results for both observers, obtained with and without a fixation dot, are tabulated in Table 1. In order to search for statistically significant differences dependent upon the presence of the fixation dot, a t-test was performed. To control for inflation of spurious significant differences in the statistical analysis, the chosen significance level of 5% has been scaled by the number of comparisons. For the fixation data, there were 4 comparisons, giving a corrected significance level of 1.25% (Subsequent statistical analyses are similarly corrected for the number of comparisons.) For both observers, the local transitions, measured with and without a fixation dot, are not significantly different at the 1.25% level. For the global transitions, results for observer JF are significantly different while those for TKD are not. For our purposes, it is particularly important that the local transitions for both observers are not significantly different with respect to the presence of a fixation dot. It suggests that the hysteresis can not be attributed to eye movements entrained to global flow.

We next examined the effects brought about by systematically altering the stimulus history. Specifically, we carried out experiments for two additional ranges of signal, 180° and 1°, keeping the noise distribution the same. Note that each of the different signal distributions will generate a different history for the directional content of the stimulus. Figure 2 shows the results for both observers at the two additional ranges, together with the original results at the 90° signal range.

From Figure 2 it can be seen that decreasing the signal range has the effect of narrowing the hysteresis profile and shifting it to the left. The leftward shift indicates that as the signal range is decreased, a smaller number of signal dots is required for a transition. This finding is not unexpected if hysteresis is indeed dependent upon the directional content of the stimulus, since a dot with a small directional range about the vertical is a more effective stimulus for upward movement than is one with a broad directional range. That is, fewer signal dots should be needed to switch from a local to a global percept when the dots chose their directions of motion from a small rather than a broad distribution. Similarly, for a smaller signal range, fewer signal dots should be required to maintain the global percept once it established. Thus, the smaller the range of the *signal* distribution, the smaller the proportion of signal dots required for the perceptual transitions. The observed shift in the hysteresis profile with decreasing signal range is further evidence that the directional content of the stimulus is a contributing factor to the hysteresis.

Spatial properties

Since the stimulus motion vectors are distributed over space, it is natural to consider how the spatial dimension figures into the cooperative behavior we have demonstrated. Also, for the purpose of formulating a mathematical model of this behavior, we need an understanding of its spatial dependence. Accordingly, we have studied the effects of changes in several spatial parameters of the display: dot step size, dot density, display area and location of the stimulus field. Effects were measured for a decrease in step size by a factor of nine, a decrease in dot density and display area by a factor of four, and a displacement of the stimulus field 8° into the periphery.

1. Step Size

Transitions were measured using a smaller step size of 0.1° for all three ranges of the signal distribution (i.e., 180, 90 and 1°). For this step size, it was necessary to decrease the interframe interval from 95.0 to 25.0 msec in order to generate smooth, continuous apparent motion.[1] These changes resulted in a decrease in dot velocity by a factor of about 3, from 8.6°/sec to 2.9°/sec. It was also necessary to decrease the rate of change in the proportion of signal, from 2 dots to 1 dot per frame, so as to maintain a comparable duration for the presentation.

Data obtained at both step sizes, 0.1° and 0.9°, are tabulated in Table 2 and shown in Figure 5. To ascertain the statistical significance of changing the step size, a t-test was carried out between respective pairs of transitions for the two conditions. The results of these tests are also included in Table 2. For observer TKD, statistically significant differences at the 0.4% level are found at only two transitions: the local transition for signal range 90° and the global transition for signal range 1°. For observer JF, all but two transitions show significant differences with a change in step size. The two that are not significantly different are the global and local transitions for the 90° signal range. In view of the differences between the observers, we determined the proportions of the variance (ω^2) associated with changing the step size. In Table 2, the proportion of the total variance that is accounted for by the change in step size is listed for each transition (see Keppel[12] for a formulation of the magnitude of treatment effect, ω^2). The largest proportion, $\omega^2 = .37$, was obtained for observer JF at the global transition for the 1° signal range. At all other transitions, the proportion of the variance that could be attributed to the change in step size was at most .16. Thus, while significant differences resulted from a change in step size, the magnitude of the effect, as a proportion of the total variance of the data, is relatively small.

2. Dot Density, Field Area and Field Eccentricity

The effects of changing the dot density, display area and eccentricity are reported together. For observer JF, measurements were made using a signal range of 90° and a step size of 0.9°. For observer TKD, the signal range was 180° and the step size, 0.1°.

The dependence of hysteresis on dot density was assessed by decreasing the number of dots by a factor of four, thus reducing the density from 1.6 to 0.4 dots/deg^2. To maintain the same rate of change in the proportion of signal and noise as with a full complement of dots, the rate of change for the 0.1° step size was reduced from 1 dot per frame to 1 dot every 4 frames, and for the 0.9° step size, from 2 dots per frame to 2 dots every 4 frames.

In order to determine the role of stimulus area, the circular display field was reduced in area by a factor of four. The effect of location was examined by centering this smaller field 8° in the nasal visual field. For both of these manipulations the dot density was maintained at 1.6 dots/deg^2, the original value, so that when the field area was reduced by a factor of four, the total number of dots presented was equal to that for the reduced density case.

The data for these experimental conditions are presented in Fig. 3 for observer JF, and in Fig. 4 for TKD. The original results for both observers at the appropriate signal range and step size are also shown in the Figures as the data sets labeled "A". Data for the reduced density stimuli are labeled "B", while data for the reduced stimulus area are labeled "C". The peripheral presentation data are labeled "D". These data are summarized in Table 3 for both observers.

To determine the statistical significance of each spatial manipulation, a t-test was performed between appropriate transition points obtained with the original display conditions and each of the other conditions. Results of these tests are presented in Table 3. For observer TKD, the data from one condition at each transition was found to be significantly different, at the 0.4% level, from the data measured using the original display parameters . These are the reduced display area data at the local transition and the peripheral presentation data at the global transition. In the case of observer JF, the peripheral presentation data were found to be significantly different from the original data at both the local and global transitions. Observer JF also showed a significant difference at the global transition for the reduced display area condition compared to results for the original display parameters. In view of the fact that the majority of the results were not significantly altered by changes in the spatial properties of the stimulus, we calculated the magnitude of the effect of each of the spatial changes. The proportion of the total variance accounted for by the manipulation of each of the spatial properties tested is listed in Table 3. For both observers, the proportion of the variance associated with each of the changes in spatial parameters is relatively small, with a maximal value of .20.

In summary, the changes in spatial properties of the display did produce some statistically significant differences in the hysteresis profiles. However, *post hoc* statistical analysis indicates that the magnitudes of such differences are small. Undoubtedly, extreme changes in the spatial properties of the stimulus would substantially alter the hysteresis characteristics but, as a first approximation, we neglect the spatial properties of the stimulus in the formulation of a cooperative model.

MODEL

As will be recalled, a cooperative system is defined as one consisting of local elements that interact to generate global behavior. The local elements in our cooperative model are a set of direction selective mechanisms. The interactions among these mechanisms consist of nonlinear excitation and inhibition such that those mechanisms with similar preferred directions of movement facilitate one another's responses, whereas those mechanisms whose preferred directions are further removed inhibit one another's responses.

Specifically, the model comprises N direction-selective mechanisms, each with a Gaussian-shaped sensitivity profile. For the k^{th} mechanism centered along direction $\varnothing_k$ the sensitivity to the direction of motion, $\varnothing$, is given by:

$$S_k(\varnothing) = A \exp \{-[(\varnothing-\varnothing_k)/h]^2 * \ln 2\} \qquad (1)$$

where h is the half-amplitude, half-bandwidth of the mechanism and A is the amplitude. These mechanisms are assumed to be evenly spaced over 360°, with adjacent mechanisms having a center-to-center separation equal to their half-amplitude half-bandwidth. The excitatory component of the k^{th} mechanism's response at time t is denoted by $E(\varnothing_k,t)$. Inhibition is mediated by a set of N associated mechanisms, with the inhibitory component of the k^{th} mechanism's response at time t given by $I(\varnothing_k,t)$.

The dynamic response of this cooperative system is represented by a pair of coupled differential equations with the form:

$$\frac{d}{dt}E_k(t) = - E_k(t) +$$

$$[1-r_e E_k(t)]\mathcal{S}_e\{\alpha\mu[\sum_{j=1}^{N}\beta_{ee}E_j(t)-\sum_{j=1}^{N}\beta_{ie}I_j(t)+\sum_{\emptyset=1}^{360}S_k(\emptyset)pr\{D(\emptyset)\}]\}$$

(2)

$$\frac{d}{dt}I_k(t) = - I_k(t) +$$

$$[1-r_i I_k(t)]\mathcal{S}_i\{\alpha\mu[\sum_{j=1}^{N}\beta_{ei}E_j(t)-\sum_{j=1}^{N}\beta_{ii}I_j(t)]\}$$

where $pr[D(\emptyset)]$ is the proportion of dots in the distribution $D(\emptyset)$ that move in direction $\emptyset$ and $\mathcal{S}_j$ is a sigmoid non-linearity of the form:

$$\mathcal{S}_j(\bar{M}_j) = [1+\exp(-\nu_j(\bar{M}_j-\theta_j))]^{-1} - [1+\exp(\nu_j\theta_j)]^{-1} \qquad (3)$$

where $j=e,i$. Interactions among the mechanisms are defined by the connectivity functions $\beta_{jj'}$ in Eq. (2). The magnitude of the interaction between a mechanism centered at $\emptyset_k$ and one centered at $\emptyset_1$ is:

$$\beta_{jj'} = b_{jj'}\exp[-|\emptyset_k-\emptyset_1|/\sigma_{jj'}] \tag{4}$$

where j=e,i and j'=e,i. The form of Eqs. (2)-(4) was originally proposed by Wilson and Cowan[13] in their cooperative theory of cortical tissue dynamics. The reader is referred to their paper for a detailed description of the parameters in Eqs. (2)-(4) and a general discussion of the model's behavior under various input conditions.

Based upon previous results obtained in our laboratory, the number of mechanisms, N, was set equal to 12 and the half-amplitude half-bandwidth, h, to 30°.[14] The parameters of Eqs. (2)-(4) were constrained in order for the system of equations to operate in what Wilson and Cowan[13] termed the active transient mode. In this mode, the system exhibits hysteresis switching between different steady states of activity. In the model simulation, the percept of local random motion is represented by a steady state of uniform activity across all mechanisms. Global upward flow is represented by a steady state in which the activity is localized about the mechanism selective for upward movement. A transition point is defined by the proportion of signal at which the network switches between these two states of activity. The results are shown in Fig. 5, with the dashed lines marking the transition points calculated from the model. For each observer, a single parameter set has been chosen to fit the data for all three different signal ranges at both step sizes. As can be seen, the model captures the leftward shift and narrowing of the hysteresis profile with decreasing signal range. This may be understood by considering that with decreasing signal range, more activity is confined to fewer direction-selective mechanisms in the neighborhood of the upward direction. Thus a smaller proportion of signal is required to indicate the upward direction of motion. Furthermore, having fewer such active mechanisms also reduces the strength of the cooperative interactions, resulting in a narrower hysteresis profile. Fender and Julesz[3] obtained a similar narrowing of binocular hysteresis profiles when the number of stimulus elements was decreased.

The model parameters for each observer are listed in Table 4. Note that the parameters for both observers differ only in a single value, specifically, the amplitude of the mechanisms' sensitivity profile. These parameter sets are not the only ones that could have been used to fit the data. However, their uniqueness is not of particular concern since we sought only to demonstrate that the hysteresis data could be interpreted in the context of a cooperative model.

DISCUSSION

We have found hysteresis in the global motion percept which results from the combination of different, localized motion vectors. Furthermore, the hysteresis characteristics are rather robust with respect to changes in the spatial parameters of the display, including dot density, display area and location, as well as step size. This relative spatial invariance suggests a form of local cooperative processing.

We did find that the hysteresis profile was sensitive to changes in the directional content of the stimulus. Specifically, narrowing the directional range of the signal brought about both a narrowing and a shift in the position of the hysteresis profile. Such behavior is consistent with cooperative processing and we have been able to describe it by a model incorporating cooperative interactions among direction-selective motion mechanisms. Both our experimental and theoretical results provide further support for a cooperative interpretation of movement perception in random-dot cinematograms, as initially proposed by Chang and Julesz.[2]

What might the role of hysteresis, and more generally cooperative processing, be in sensory processing? By the very nature of its interactions, a cooperative network is well-suited for the enhancement of signal in a noisy environment.[15] In the case of binocular hysteresis, cooperative processing will make the ocular registration necessary for binocular stereopsis relatively resistant to noise. With respect to motion perception, the function of cooperative interactions among direction-selective mechanisms may be to enhance the perception of unidirectional flow in the midst of noise.

References

1. D. Williams and R. Sekuler, "Coherent global motion percepts from stochastic local motions," Vision Res. 24, 55-62 (1984).

2. J.J. Chang and B. Julesz, "Co-operative phenomena in apparent movement perception of random-dot cinematograms," Vision Res. 24, 1781-1788 (1984).

3. D. Fender and B. Julesz, "Extension of Panum's fusional area in binocularly stabilized vision," J. Opt. Soc. Am. 57, 819-830 (1967).

4. G. Sperling, "Binocular vision: a physical and a neural theory," Am. J. Psychol. 83, 461-534 (1970).

5. P. Dev, "Perception of depth surfaces in random-dot stereograms: a neural model," Int. J. Man-Mach. Stud. 7, 511-528 (1975).

6. J.J. Nelson, "Globality and stereoscopic fusion in binocular vision," J. Theoret. Biol. 49, 1-88 (1975).

7. D. Marr and T. Poggio, "Cooperative computation of stereo disparity," Science 194, 283-287 (1976).

8. J.E.W. Mayhew, J.P. Frisby and P.Gale, "Computation of stereo disparity from rivalrous texture stereograms," Perception 6, 207-208 (1977).

9. H.R. Wilson, "Hysteresis in binocular grating perception: contrast effects," Vision Res. 17, 843-851 (1977).

10. P. Tynan and R.W.Sekuler, "Rapid measurement of contrast sensitivity functions," Am. J. Optom. Physiol. Opt. 54, 573-575 (1977).

11. W. Kosnik, J. Fikre and R. Sekuler, "Improvement in direction discrimination: No role for eye movements," Percept. and Psychophy. 38, 554-558 (1985).

12. G. Keppel, Design and Analysis: A Researcher's Handbook (Prentice-Hall, Englewood Cliffs, N.J., 1973).

13. H.R. Wilson and J.D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," Kybernetik 13, 55-80 (1973).

14. D. Williams, S. Tweten and R. Sekuler, "Using motion metamers to investigate the mechanisms of motion," Supplement to Invest. Ophthal. and Vis. Sci. 25, 14 (1984).

15. H.R. Wilson, "Cooperative phenomena in a homogeneous cortical tissue model," in Synergetics, H. Haken, ed. (B.G. Teubner, Stuttgart, 1973).

TABLE 1

Comparison of Transition Data With and Without Fixation Mark

OBSERVER	TRANSITION	WITHOUT FIXATION MARK	WITH FIXATION MARK	t-STATISTIC $P<0.0125$
JF	local	$.228 \pm .113$	$.235 \pm .093$	$t_{177}=0.471$
	global	$.553 \pm .073$	$.597 \pm .099$	$t_{177}=3.327$*
TKD	local	$.347 \pm .112$	$.314 \pm .107$	$t_{118}=1.674$
	global	$.683 \pm .152$	$.698 \pm .152$	$t_{118}=0.654$

* = statistically significant

TABLE 2

Comparison of Transition Data for Two Different Step Sizes

BSERVER	SIGNAL RANGE	TRANSITION	STEP SIZE 0.9°	0.1°	t-STATISTIC P<.004	MEASURE OF EFFECT MAGNITUDE (ω^2)
JF	180°	local	$.264 \pm .159$	$.166 \pm .148$	$t_{195}=4.461$*	0.09
		global	$.776 \pm .122$	$.842 \pm .160$	$t_{195}=3.279$*	0.05
	90°	local	$.228 \pm .113$	$.211 \pm .116$	$t_{178}=0.956$	<0.01
		global	$.553 \pm .073$	$.597 \pm .126$	$t_{178}=2.787$	0.04
	1°	local	$.051 \pm .062$	$.085 \pm .082$	$t_{197}=3.251$*	0.05
		global	$.351 \pm .059$	$.462 \pm .083$	$t_{197}=10.89$*	0.37
TKD	180°	local	$.378 \pm .160$	$.321 \pm .157$	$t_{198}=2.521$	0.03
		global	$.808 \pm .159$	$.818 \pm .138$	$t_{198}=0.446$	<0.01
	90°	local	$.347 \pm .112$	$.250 \pm .103$	$t_{158}=5.590$*	0.16
		global	$.683 \pm .101$	$.717 \pm .099$	$t_{158}=2.090$	0.02
	1°	local	$.139 \pm .096$	$.125 \pm .093$	$t_{197}=1.066$	<0.01
		global	$.343 \pm .076$	$.395 \pm .099$	$t_{197}=4.091$*	0.07

* = statistically significant

TABLE 3

Comparison of Transition Data
For Four Different Display Conditions

OBSERVER	TRANSITION	DISPLAY CONDITION[1]	TRANSITION VALUE	t-STATISTIC P<0.004	MEASURE OF EFFECT MAGNITUDE (ω^2)
JF	local	A	$.228 \pm .113$		
		B	$.241 \pm .110$	$t_{178}=0.786$	<0.01
		C	$.186 \pm .149$	$t_{178}=2.077$	0.02
		D	$.171 \pm .137$	$t_{178}=2.966$*	0.04
	global	A	$.553 \pm .073$		
		B	$.572 \pm .100$	$t_{178}=1.448$	<0.01
		C	$.634 \pm .121$	$t_{178}=5.261$*	0.13
		D	$.657 \pm .123$	$t_{178}=6.690$*	0.20
TKD	local	A	$.321 \pm .157$		
		B	$.314 \pm .147$	$t_{158}=0.312$	<0.01
		C	$.226 \pm .134$	$t_{198}=4.646$*	0.09
		D	$.292 \pm .156$	$t_{158}=1.163$	<0.01
	global	A	$.818 \pm .138$		
		B	$.871 \pm .108$	$t_{158}=2.573$	0.04
		C	$.865 \pm .097$	$t_{198}=2.784$	0.03
		D	$.737 \pm .075$	$t_{158}=6.584$*	0.06

* = statistically significant
[1] for explanation of letters see caption of Figure 3

TABLE 4

Model Parameter Values {Equations (1)-(4)}

PARAMETERS	VALUES
N	12.0
h	30.0°
α	1.0
μ	10.0
r_e	1.0
r_i	1.0
ν_e	0.5
θ_e	9.0
ν_i	0.174
θ_i	8.0
b_{ee}	25.5
σ_{ee}	77.0
b_{ie}	22.95
σ_{ie}	115.5
b_{ei}	22.95
σ_{ei}	115.5
b_{ii}	30.6
σ_{ii}	57.8
A{	28.88 (observer JF) 31.75 (observer TKD)

FIGURE CAPTIONS

Figure 1. Data from two observers (JF,TKD) showing the transitions in the percept of motion direction for two different histories of stimulus exposure (shown by arrows). The solid circles indicate the proportion of "signal" dots required for the transition from random, local motion to global, upward flow (G); the open circles indicate the proportion required for the transition from global, upward flow to random, local motion (L). Error bars represent one standard deviation. The range of the signal distribution was 90°. The separation between transition points within each panel is a measure of hysteresis. Step size, 0.9°.

Figure 2. Hysteresis profiles from the same observers at signal ranges of 180°, 90° and 1°. Note the narrowing and leftward shift of the profiles with decreasing signal range. Step size, 0.9°.

Figure 3. Comparison of results from observer JF for A) original display parameters, B) four-fold decrease in dot density, C) four-fold decrease in display area and D) four-fold decrease in display area plus displacement of field 8° into nasal periphery. The open symbols represent "local" transition data; the closed

symbols, "global" transition data. The signal range is 90° for a 0.9° step size.

Figure 4. As in Figure 3, for observer TKD with a signal range of 180° and a step size of 0.1°.

Figure 5. The format here is as in Figure 2, but with additional data for the 0.1° step size indicated by square symbols. The dashed lines mark the transition points calculated from a cooperative model (see text). Note that the model captures both the leftward shift and narrowing of the hysteresis profile with decreasing signal range.

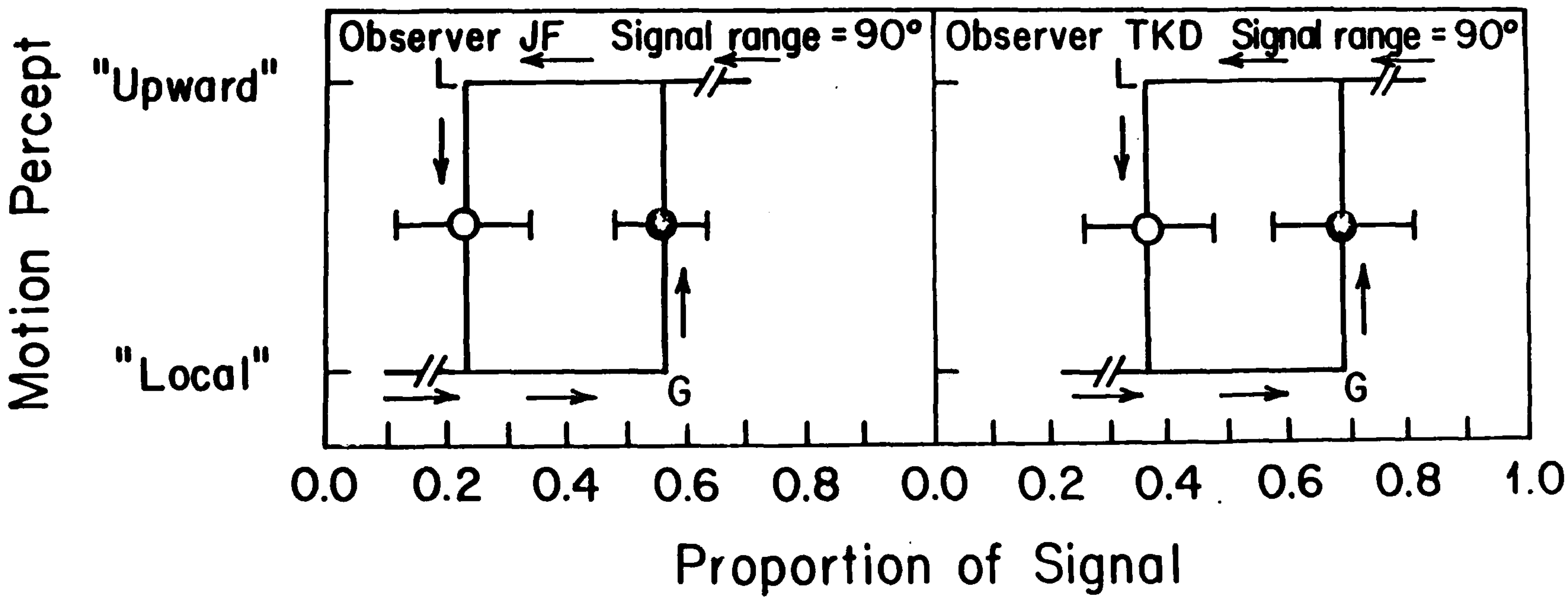

Observer JF
Signal range = 90°
Observer TKD
Signal range = 90°
L
G
"Upward"
"Local"
Motion Percept
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of Signal

Figure 1.

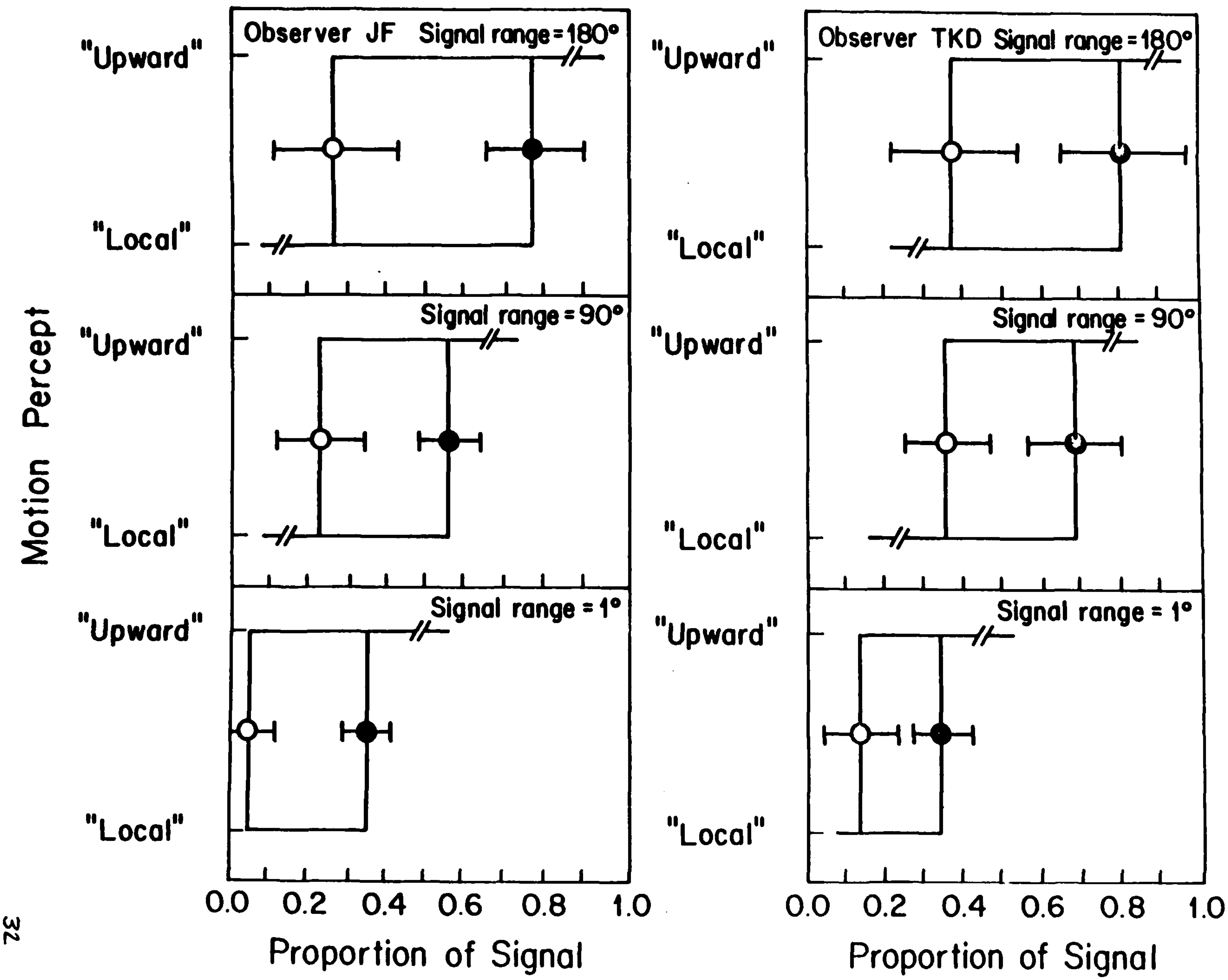

Observer JF
Signal range = 180°
Signal range = 90°
Signal range = 1°
Observer TKD
Signal range = 180°
Signal range = 90°
Signal range = 1°
Motion Percept
"Upward"
"Local"
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of Signal

Fig 2

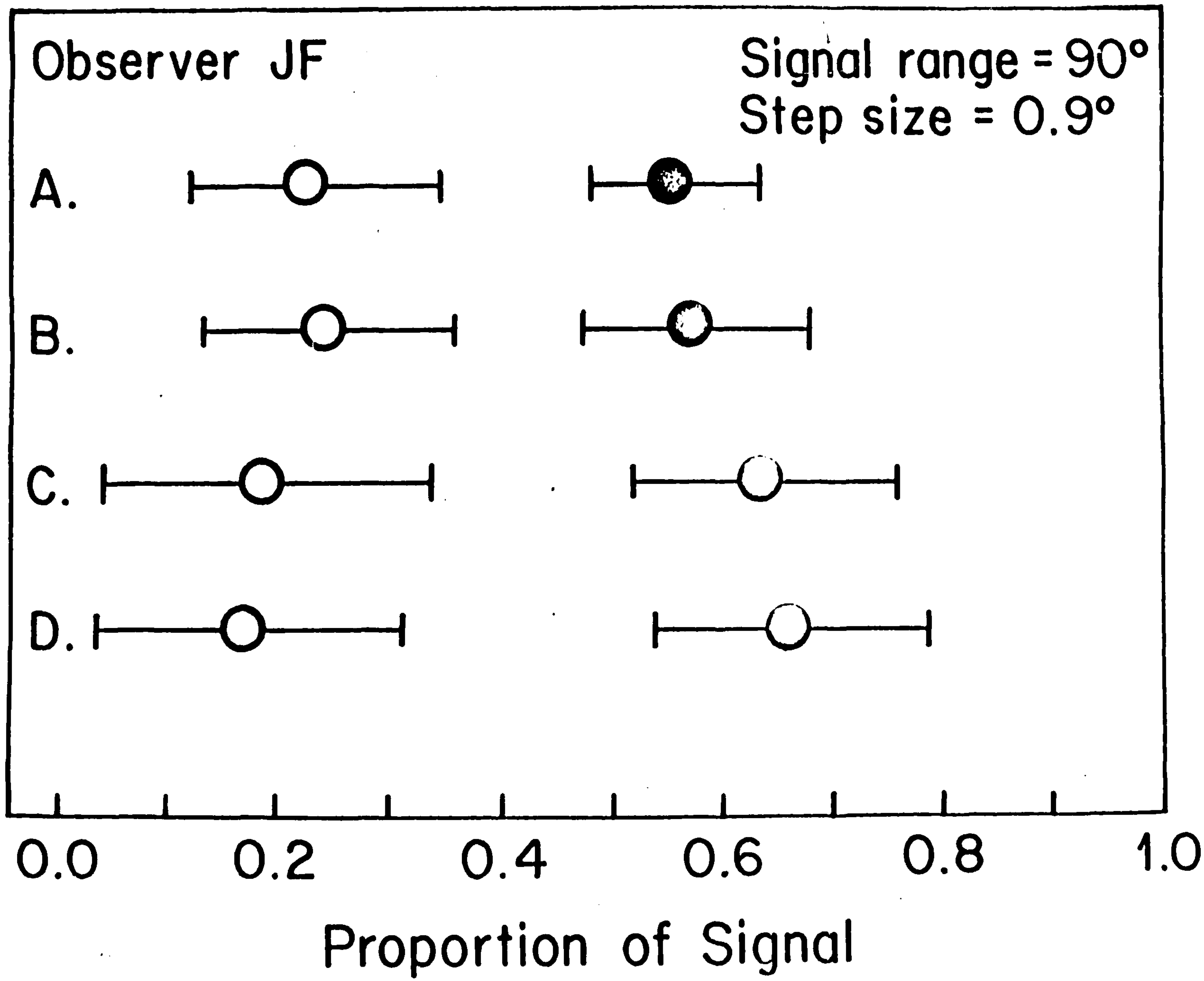
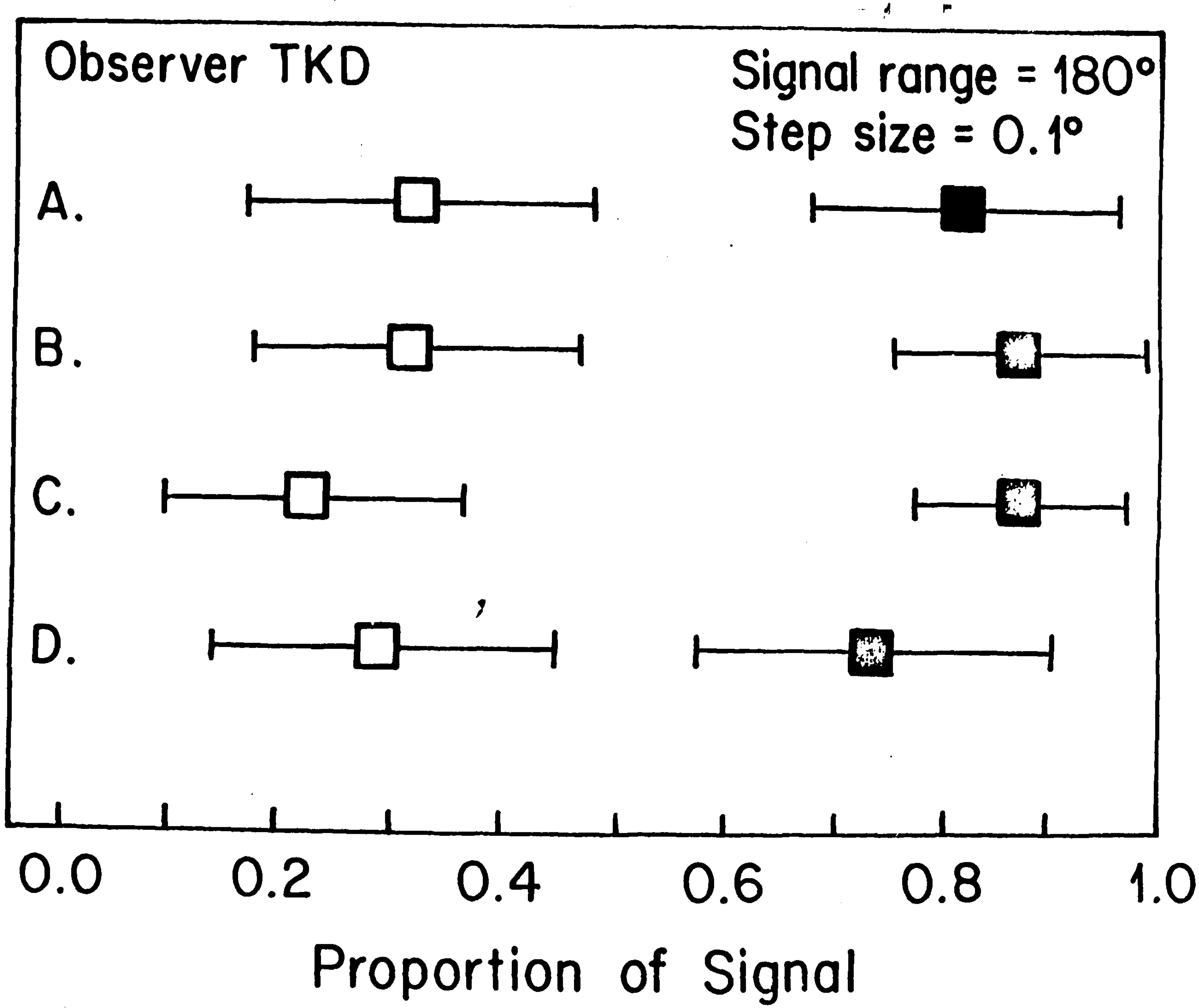

Observer JF
Signal range = 90°
Step size = 0.9°
A.
B.
C.
D.
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of Signal

Figure 3


Observer TKD
Signal range = 180°
Step size = 0.1°
A.
B.
C.
D.
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of Signal

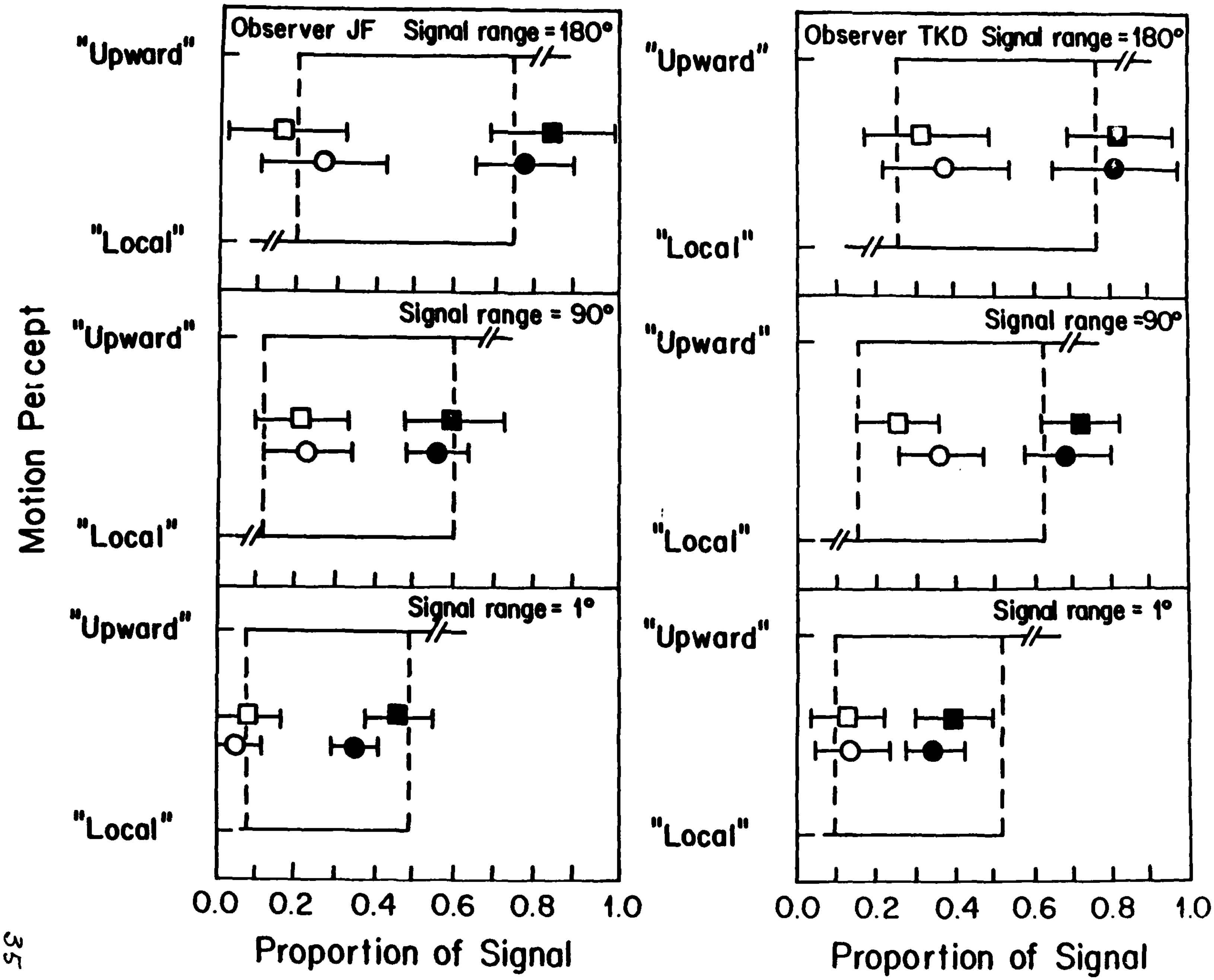

Observer JF
Signal range = 180°
Signal range = 90°
Signal range = 1°
Observer TKD
Signal range = 180°
Signal range =90°
Signal range = 1°
"Upward"
"Local"
Motion Percept
0.0 0.2 0.4 0.6 0.8 1.0
Proportion of Signal

Figure 5

Project Two:

Role of eye movements in improving direction discrimination

William Kosnik, John Fikre,
and Robert Sekuler

Practice improves an observer's ability to discriminate one direction of movement from another highly similar direction of movement (Ball and Sekuler, 1982). This improvement in discrimination has two noteworthy features, directional selectivity and persistence. More particularly, the improvement is restricted to directions that are similar to the one with which the observer has practiced, and the improvement endures for several months without noticeable decrement. We sought to clarify the origin of this direction-specific change in discrimination.

Basically, improved direction-discrimination could be achieved through two different routes. For one, the route may be purely visual, possibly reflecting changes in the selectivity of neurons at some stage of the visual system. Alternatively, the route may be sensori-motor, with the observer learning to use tracking eye movements to discriminate between two directions.

In support of this second possibility McHugh and Bahill (1985) have shown that an observer can learn to use smooth pursuit movements to track a novel target and that the movements are specific to the waveform of the target. They have also shown that, once learned, the observer retains this ability over a long period of time. Given this ability of the oculo-motor system, we sought to determine if eye movements play a role in an observer's lea-ning to discriminate the direction of moving targets.

In their original paper Ball and Sekuler (1982) did measure the eye movements of two observers and found steady fixation with high levels of performance. However, significant questions about the role eye movements remained unanswered. Because their recording system could not resolve eye movements smaller than approximately 45 minutes of arc, Ball and Sekuler were unable to rule out the possibility of small, but visually significant, eye movements. This possibility gains importance since the stimulus duration they used, 500 msec, might prevent very large pursuit movements anyway. More importantly, though, they neglected to record eye movements at different stages of training. Therefore, it remains possible that <u>changes</u> in eye movements might have played some role in the observed change of performance.

We decided to investigate sensori-motor contributions to direction discrimination more thoroughly by analyzing an observer's eye movements at the beginning and end of training using an eye tracking device that is capable of resolving movements of about one minute of arc.

METHOD

<u>Observer</u>

The observer was a 20-year old male who had never participated in a psychophysical study before. He was paid $7.50/hour for his participation. Also, to insure high motivation, he received an additional one cent for every correct response. This was the same motivational device used in the earlier work on motion discrimination (Ball and Sekuler, 1982).

The observer viewed the stimulus display with the right eye;

the other eye was occluded with an opaque patch.

Apparatus

The experimental set up was similar to that used by Ball and Sekuler. The stimuli were 512 spatially-random dots moving in a uniform direction at 10 degrees per second across the face of a cathode ray tube (CRT). The dots were plotted under computer control at a framerate of 28.5 Hz, again similar to that used in the earlier study. The dots, which appeared within a circular aperture of 5 degrees, had a luminance of 104 cd/m^2. They were easily visible against the CRT's luminance of 2.06 cd/m^2. A small fixation point was provided in the center of the screen.

Procedure

A trial consisted of two stimulus presentations, each lasting 640 msec (except on the first day of training, when each presentation lasted 512 msec). The two presentations were separated by a interval of 1.25 sec, during which the CRT was blank.

The directions of movement of the dots within the two presentations were either the Same --in both presentations the dots moved in a direction of 90 degrees from horizontal (upward)-- or Different --during one interval the dots moved upward and in the other interval the dots moved either 3 degrees to the left (93°) or 3 degrees to the right (87°) of upward. Same and Different trials were randomly presented with equal probability. For Different trials the computer randomized whether the upward movement would occur in the first interval or in the second. Also for Different trials the two non-upward directions,

87° and 93°, occurred randomly, but equally often.

After each trial the observer judged whether the two directions had been the same or different, that is, whether both stimuli moved in the upward direction or whether one moved in the upward direction and the other moved in a direction other than upward. A computer-generated tone provided knowledge of the correctness of the observer's response.

Training comprised an extended series of discrimination trials in blocks of 32 trials each. Because half the trials were <u>Same</u> and half <u>Different</u> and because there were two stimulus presentations per trial, every block of 32 trials yielded 64 stimulus presentation intervals --48 in which movement was upward, 8 in which movement was in a direction of 87° degrees, and another 8 in which movement was in a direction of 93° degrees.

On the first day of training four blocks of 32 trials were run. On subsequent days ten blocks of 32 trials each were run. A rest was given after each block. Training was spread out over eight days.

<u>Eye movement recording</u>

Two dimensional eye movements were measured from the observer's right eye by an Scientific Research International (SRI) dual Purkinje Image Eye Tracker (Mark IV). This electro-optical instrument determines the instantaneous position of the eye from two reflections of a narrow infrared beam projected into the eye. One reflection originates from the anterior surface of the cornea (the first Purkinje image) and the other from the posterior surface of the lens (the fourth Purkinje image). Rotational eye movements are derived from the difference in the relative position of these two images.

The Eye Tracker's noise level was determined by tracking a stationary, artificial eye. Expressed as the standard deviation of the sampled positions of the stationary artificial eye, the Eye Tracker's noise level was 0.43 minutes of arc in the horizontal channel and 0.40 minutes of arc in the vertical channel.

The gain factors for the instrument's horizontal and vertical channels were determined by a calibration procedure in which the observer fixated a target on the CRT. This target made five steps first along the horizontal axis, and then the vertical axis, in increments of 0.25 degrees. At each increment, when the observer was satisfied that he had achieved good fixation of the target, he pressed a switch, triggering a 640 msec period of data collection. The target then moved to its next position. This procedure continued until eye positions had been recorded in response to five stimulus positions along the horizontal axis and five along the vertical axis.

After the calibration procedure, we fit a least squares regression line to the recorded eye positions that were plotted against the corresponding stimulus positions. Separate regression lines were fit to horizontal eye positions and to vertical eye positions. Horizontal and vertical gains were obtained from the regression coefficients of those regression lines. We estimated the accuracy of fixation from the correlation between the target positions and the eye positions. This correlation coefficient was at least 0.99 for each axis.

Eye position records obtained from four blocks of trials on the first day of discrimination training were digitized at a rate of 500 Hz and stored in computer memory. A 500-Hz sampling rate was used in order to accommodate the full, 200-Hz bandwidth of the recording instrument. Eye positions were collected throughout the 512 msec stimulus presentation. This yielded one eye position record of 256 data points.

On subsequent training days a 640 msec stimulus presentation interval was used. This change was necessitated by the introduction of a low pass filter in the data collection system on the second training day, as explained below.

On the last day of training, we measured eye positions during the last five blocks of discrimination training. The data were low pass filtered at 50 Hz (-36 dB/octave) prior to being digitized at a rate of 100 Hz. These recording parameters required a 640 ms stimulus presentation, but resulted in a considerable savings in computer storage without loss of significant eye position information. Thus, each eye position record collected on the last day of training contained 64 data samples.

Editing eye position records. A continuous record of the status of the Eye Tracker was obtained at the same time that an eye position was being recorded. This record contained information about the occurrence of eye blinks and occasional interruptions of tracking. If an eye blink occurred or if tracking had been interrupted at any time during a record, the entire record was omitted from the analysis. In addition, since we wanted to know whether the observer was tracking the moving stimulus, we eliminated records that contained saccades. In particular, any record containing a saccade with a velocity greater than 30° per second was omitted from the analysis.

Results

Discriminability

The observer's discrimination performance for each block of trials was expressed in units of d' (Swets, 1964), computed from the proportion of Different trials correctly identified as "different" (that is, hits) and the proportion of Same trials incorrectly identified as "different" (that is, false alarms). A discriminability score for one day was obtained by averaging across all blocks of trials run on that day.

The first question that needs to be answered is whether the observer's discrimination performance changed with practice, and, if it did, whether such changes mirror those previously reported by Ball and Sekuler. To answer these questions we have portrayed in Figure 1 the observer's discrimination performance over the eight days of training. To facilitate comparison we have plotted on the same axes the results of Ball and Sekuler (1982), which represents the average performance of eight observers. Note the similarity of the two curves, each demonstrating a steady improvement in performance and reaching the same high level of discrimination.

Eye Movements: Orientation

Next we wanted to determine if the improvement in discrimination was mediated by the observer's having learned to track the stimulus. Since tracking eye movements would cause successive samples of eye position to lie along a straight line, we developed an estimate of the main axis along which the eyes moved during each stimulus presentation. We called this estimate the dominant orientation. To obtain this dominant orientation, the eye positions recorded during an presentation interval were represented in two dimensions and a least squares regression line fit thereto. The slope of this line, expressed in degrees from the 0° meridian, defined the dominant orientation of the eye movements.

To illustrate this procedure two eye position records are shown in Figure 2; the dominant orientations have been drawn through the sampled eye positions. For each record, the *F* ratio associated with the regression coefficient is highly significant: $F=898.6$ and $F=110.4$, for the top and bottom panels respectively, both $df=1,253$ and $p<0.0001$.1

Table 1 gives the mean dominant orientation of the eye positions recorded during the first and last days of training. These dominant orientations have been sorted according to stimulus direction, with the 87° direction in the first column, the 90° direction in the second, and the 93° direction in the third. Measurements made at the beginning of training are represented in the top row and measurements from the end of training are represented in the middle row. Note that orientations are expressed as axial values, meaning that 0° and 180° are equivalent to one another. Orientations take on values from 0 to 179°. Means and variances are computed using statistics for directional data (Mardia, 1972). Standard deviations are shown in parentheses, with the number of trials included in the average shown in brackets.

As can be seen in the top row of Table 1, the mean dominant orientation for records at the beginning of training was centered near the horizontal axis (0-180°) for all three stimulus directions. Also note that there is no correspondence between the change in direction of the stimulus movement and the dominant orientation. A change in stimulus direction from 90° to 87° --a shift of 3° to the right-- is not accompanied by a corresponding change in the dominant orientation. Instead, the dominant orientation shifted from 168° to 176°, a net change of 8° to the left. A change in stimulus direction from 90° to 93° (a shift of 3° leftward) also failed to elicit a corresponding change in dominant orientation. Here, the dominant orientation shifted 6° to the right.

An examination of the middle row of Table 1 shows no better correspondence between stimulus direction and dominant orientation at the end of training. Again, the mean dominant orientation at the end of training was close to the horizontal axis for all three stimulus directions. In response to stimulus movement of 90° the dominant orientation is 24°. A change in stimulus direction of 3° to the left or right of vertical was not followed by a similar change in dominant orientation. In fact, the mean dominant orientation was 17° for both off-vertical stimulus directions.

So that the reader can better appreciate the variability in the obtained dominant orientations, Figure 3 shows the distribution of the dominant orientations cumulated over presentation intervals. These are the distributions that Table 1 summarized. The upper portion of Figure 3 portrays data collected at the beginning of training and its middle portion portrays data from the end of training. Each column represents one direction of stimulus movement: 87°, 90°, or 93°.

Note that for neither the beginning nor the end of training is there any obvious systematic relation between the dominant orientations and the direction of the stimulus motion. Moreover, there is no systematic change in the distribution of dominant orientations from beginning to end of training.

Eye Movements: Magnitude

Having characterized the dominant orientations of the eye position records, we then wanted to determine the linear distances the eye travelled along the dominant orientations. The magnitude of the dominant orientation was measured along the length of the regression line. The limits of the regression line were determined by finding the maximum and minimum values of one of the coordinates --either x or y-- and then computing the othe. coordinate from the regression equation. The distance between these two pairs of coordinates defined the magnitude of the dominant orientation of the eye position record. The lines of best fit in Figure 2 have been drawn to correspond with this definition.

We used this measure of eye movement magnitude, rather than the total distance the eye moved during a stimulus presentation, because, within any one stimulus presentation, the eye often moved in several different directions as well as back and forth along the same direction. Since we were mainly concerned with eye movements used to track the stimulus, we wanted a magnitude measure that would characterize the linear distance the eye would have moved to track a stimulus moving in a single direction. The length of the regression line defined by the limits of the eye position record best estimates this distance.

Table 2 lists the mean magnitude of the dominant orientation for all eye position records from a given day of training. The top row of the table lists the magnitudes from the beginning of training and the middle row from the end of training. Columns represent different stimulus directions. At both the beginning and end of training just one minute of arc separates the dominant magnitude associated with the three stimulus directions. Averaging across the three stimulus directions, less than two minutes of arc distinguishes the mean dominant magnitude at the beginning of training from the comparable value at the end of training.

Note that across all stimulus directions and across days of training the mean dominant magnitude was very much smaller than the distance travelled by the stimulus on either the first or last day of training --5.1° on the first day and 6.4° on the last day. The distributions of magnitudes associated with each stimulus direction are shown in the top and middle portions of Figure 3. These magnitudes range from 5.4 to 30.0 minutes of arc.

Supplementary Measures

To further characterize the eye movements made during discrimination training we measured eye movements under two additional conditions. In the first condition the observer was instructed to track the moving stimulus. In the second condition eye movements were recorded while the observer simply fixated a stationary fixation point with the stimulus absent.

Eye Movements: Intentional Tracking. The mean dominant orientations measured during intentional tracking are shown in the bottom row of Table 1. These orientations are very similar to the stimulus directions. Tracking eye movements to the 90° and 93° stimulus directions deviated on average just 2° from those directions. In response to the 87° stimulus the dominant orientation was 79°, indicating an error in tracking of 8° to the right. Nevertheless, the directions of the tracking eye movements were in the correct relation to the direction of the stimuli.

The distributions of the tracking dominant orientations for each stimulus direction are shown graphically in the bottom row of Figure 3. It can be seen that the dominant orientations cluster near the direction that the target moved. Also, note the narrow distribution of the tracking eye movements for each of the three stimulus directions.

The bottom row of Table 2 shows the average magnitude of the eye movement records taken while the subject attempted to track the stimulus. When the observer attempts to track the target, which moves 6.4 degrees, his eye moves a mean distance of 1.75 degrees. This tracking distance is nearly ten times greater than 0.18 degrees, the mean magnitude of the dominant orientation during discrimination training in which the observer was not instructed to track.

The distributions of the magnitudes of the tracking eye movements for each stimulus direction are illustrated in the bottom portion of Figure 3. The size of these movements ranged from 17 to 169 minutes of arc.

Eye Movements: Fixation. We then recorded eye positions while the observer was fixating a stationary target with no dots present. We compared these records to ones obtained under conditions of discrimination training, in which both moving dots and a stationary fixation target were present.

In the absence of moving dots, the observer's mean dominant orientation is 159° (SD = 26.8). The magnitude of the dominant component during fixation is 13 minutes of arc (SD = 7.45). These values are similar to the measurements obtained during training (see first two rows of Table 1). So the observer maintained approximately the same degree of fixation either while fixating a point on an otherwise blank screen, or while fixating the same point superimposed on a field of moving dots.

Discussion

Improvement in the discriminability of the direction in which targets move does not depend on the observer learning to track the moving target. For one thing, the eye movements recorded during training bore little resemblance to eye movements obtained when the observer deliberately tracked the stimulus. Neither the orientation nor the magnitude of the dominant linear component extracted from the eye position records matched the direction or distance the stimulus travelled. Dominant orientations were closer to the horizontal axis than the vertical axis, along which the stimulus moved. There was also considerable variability in the dominant orientation of the eye position records. The magnitudes of the dominant linear component of the eye position records were about 32 times smaller than the extent of the stimulus movement.

Also, the size and dominant orientation of eye movements were unchanged from the beginning to the end of training, although discriminability changed dramatically. In fact, both at the beginning and the end of training eye movements closely resembled fixation eye movements in magnitude and orientation.

In contrast to the lack of tracking eye movements during training, the observer was clearly able to track the stimulus when asked to do so. Here, the direction of tracking eye movements closely approximated the direction of the stimulus movement and the size of the tracking movements were about 10 times larger than the average magnitude of the dominant orientation of the eye position records during training.

Because the mean magnitude of tracking movements was only 1.75°, we were curious to discover why the tracking eye movements were smaller than the 6.4 degrees travelled by the stimulus. Three factors may be help to explain this difference. First and most important, it was clear from the tracking records that the observer did not track at the same rate of the stimulus movement. The observer tracked at a rate of about 6°/sec instead of the 10°/sec rate of the stimulus. Since the stimulus was a display of moving dots that continually filled the screen, the observer could follow the direction of the moving dots without having to fixate on a single dot. Thus, the distance covered while tracking the display could be less than the distance covered while tracking a single dot. Second, although the dots moved a total of 6.4°, the diameter of the viewing aperture was only 5°. Thus, the maximum distance the observer could track the stimulus would be only 5°. Third, at the start of testing, the observer reacted to the onset of stimulus movement with an appreciable latency, about 200 msec. Such a delay in the start of tracking would shorten the total tracking distance. Taken together, these three factors account for the shorter mean distance of the observer's tracking eye movements compared to the distance the target moved.

We found that after some practice at tracking, the observer managed to reduce the latency of his tracking response to as little as 10 msec. This finding is consistent with the report of McHugh and Bahill (1985) who found that observers were able to learn to track a target that had a predictable onset with no delay.

Finally, our results have answered the question with which we began: improvement in direction discrimination with practice is the product of a change in a visual process, rather than some change in sensori-motor response. With this clarification in hand, research can now attempt to delineate the visual processes that give rise to long-lasting, direction-specific improvement in discrimination.

REFERENCES

Ball, K., and Sekuler, R. (1982) A specific and enduring improvement in visual motion discrimination. Science 218 697-698.

McHugh, D.E., and Bahill, A.T. (1985) Learning to track predictable target waveforms without a time delay. Investigative Ophthalmology & Visual Science 26 932-937.

Mardia, K.V. (1972) Statistics of Directional Data. New York: Academic Press.

Swets, J.A. (1964) Signal Detection and Recognition by Human Observers. New York: Wiley.

Footnote

1. Note that our procedure for estimating dominant orientation of the eye positions assumes that the eye's excursions can be described by a linear function. Although a record may be associated with a significant regression coefficient, it does not imply that it can be completely described by a linear model. A test of the lack of fit to a linear model shows a significant departure from linearity in the top panel ($F=2.35$, $df=39,214$, $p<.001$), but not in the bottom panel ($F=0.91$, $df=23,230$). The record in the top panel departs from linearity because it contains other non-linear components. Since our main concern is to discover if the eye moved in the same direction as the stimulus, by assuming that each record contains a significant linear component, it would be possible to find out if the orientation of this linear component matches the direction along which the stimulus traveled.

Table I

Mean Dominant Axial Orientations of the Eye Position Records (in Degrees)

Training Session	Stimulus Direction		
	87°	90°	93°
Beginning of Training	176 (18.5) [28]	168 (20.3) [172]	162 (19.8) [27]
End of Training	17 (26.5) [40]	24 (27.3) [235]	17 (19.8) [40]
Intentional Tracking	79 (2.67)	88 (2.39)	91 (1.30)

Note-Standard deviations are shown in parentheses; the number of orientations included in each mean is shown in brackets.

TABLE II

Mean Magnitudes of the Dominant Axial Orientations of the Eye Position Records (in Minutes of Arc)

Training Session	Stimulus Direction		
	87°	90°	93°
Beginning of Training	12 (3.66) [28]	11 (3.18) [172]	11 (2.64) [27]
End of Training	11 (7.80) [40]	10 (4.32) [235]	10 (2.82) [40]
Intentional Tracking	96 (30.1)	100 (37.9)	122 (36.2)

Note-Standard deviations are shown in parentheses; the number of orientations included in each mean is shown in brackets.

Figure Captions

Figure 1. Discriminability (d') of the direction of a moving target as a function of the number of training days. The figure compares the performance of the observer in this study (B.McH.) with the average performance of eight observers in the Ball & Sekuler (1982) study.

Figure 2. A two dimensional eye position record collected during one stimulus presentation is displayed in each panel. A least squares regression line is fit to each record and represents the dominant axial orientation of the eye position record. The length of the regression line defines the hypothetical distance the eye moved along its dominant orientation. The calculation of this distance is described in the text. Note that, although the two records are well described by a straight line, the record in the top panel departs significantly from a linear model whereas the record in the bottom panel is completely described by a linear model (see footnote 1).

Figure 3. Distributions of the dominant axial orientations of the eye position records arranged according to recording session and direction of stimulus movement. The figure also displays the magnitudes of the orientations. Note that the beginning and end of training magnitude records are plotted on a scale of 60 minutes of arc; the magnitudes recorded during

intentional tracking of the stimulus are plotted on a scale of 300 minutes of arc.

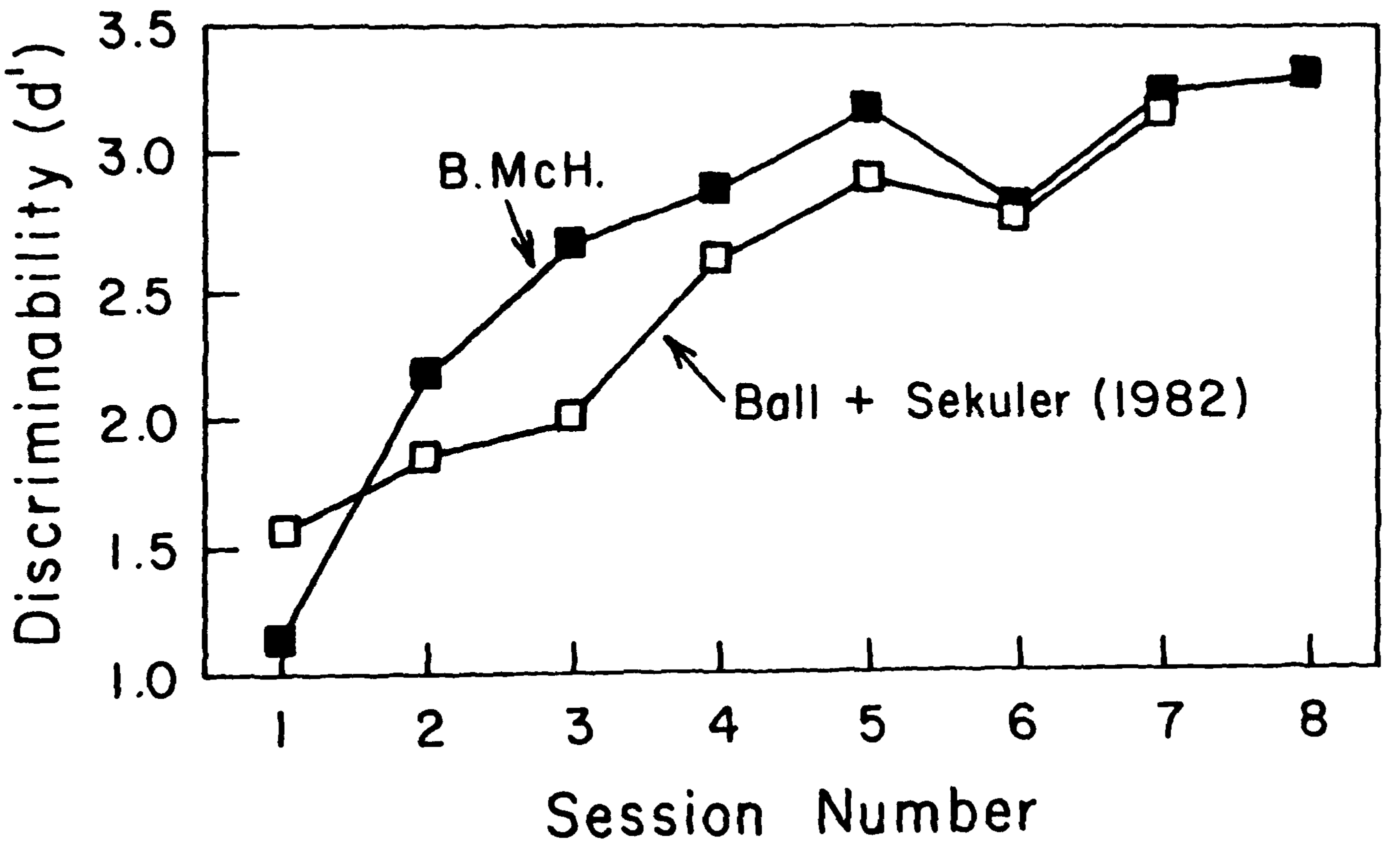
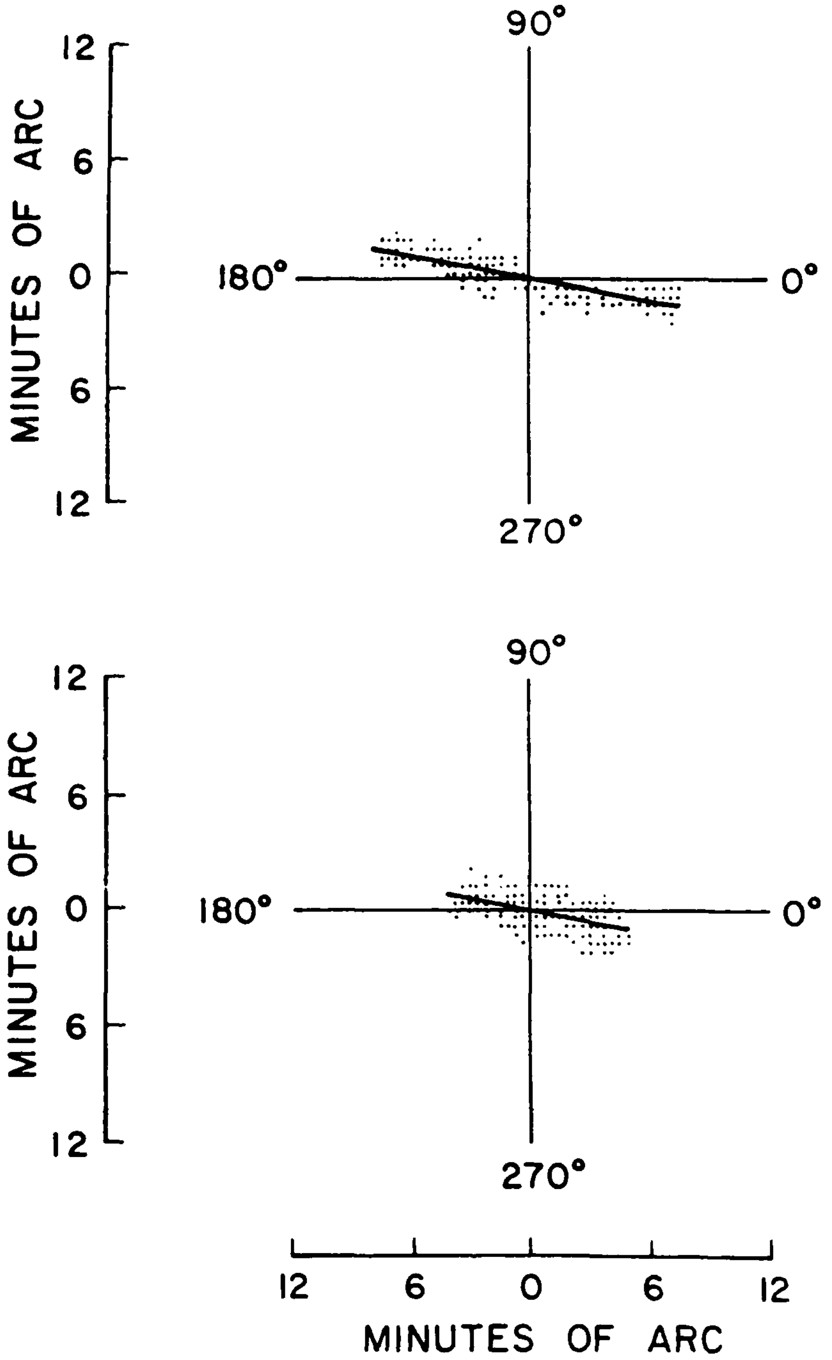

Discriminability (d')
3.5
3.0
2.5
2.0
1.5
1.0
B.McH.
Ball + Sekuler (1982)
1
2
3
4
5
6
7
8
Session Number

Figure 1



90°
12
6
0
6
12
MINUTES OF ARC
180°
0°
270°
90°
12
6
0
6
12
MINUTES OF ARC
180°
0°
270°
12 6 0 6 12
MINUTES OF ARC

Figure 2

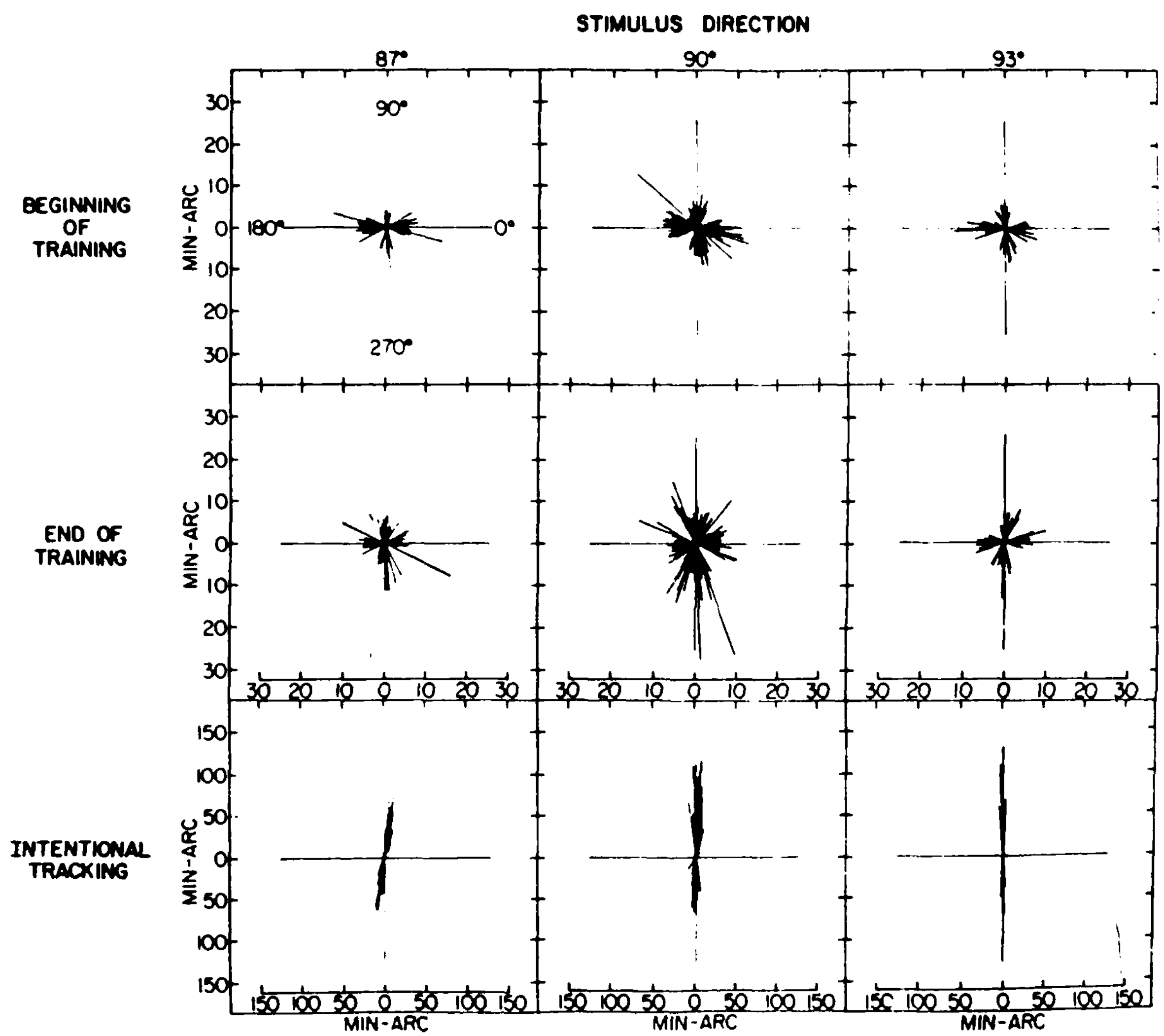

STIMULUS DIRECTION
87°
90°
93°
BEGINNING OF TRAINING
END OF TRAINING
INTENTIONAL TRACKING
MIN-ARC
90°
180°
0°
270°
30 20 10 0 10 20 30
150 100 50 0 50 100 150

Figure 3

Project Three:

Direction Perception in Complex Dynamic Displays: The Integration of Direction Information

Scott Watamaniuk, Robert Sekuler,
and Douglas W. Williams

INTRODUCTION

Though motion perception does depend upon spatially local processes, under certain circumstances global processes make an important contribution. For example, the human visual system can integrate different, spatially-intermingled motion vectors into a global percept of motion in a single direction (Adelson and Movshon, 1982; Williams and Sekuler, 1984). Such integrated percepts may offer important clues to the mechanisms of motion perception. To exploit such clues we have followed the tradition of using discrimination performance to probe underlying psychophysical mechanisms (e.g., Graham, 1965; Wilson and Gelb, 1984). Specifically, we were interested in how easily observers could discriminate between two different global motions when each had resulted from the integration of many different motion vectors.

Our stimuli were random dot cinematograms in which each dot took an independent two-dimensional random walk with steps of constant size. The direction any dot moved, from one display frame to the next, was independent of the dot's previous movements as well as the movements of other dots. All dots chose their directions of movement from the same probability distribution. Williams and Sekuler (1984), using uniform distributions of directions, showed that the resulting global percept of motion depends upon the range of the distribution. Specifically, uniform distributions with ranges of directions less than 180° tend to produce a perception of global motion in the approximate direction of the distribution's mean even though the random perturbations of each dot are evident. As the range increases further, the perception of global motion diminishes, until at the limit, a uniform distribution with 360° yields a percept of only local random motion of individual dots. In this present study, we measured the discriminability of the direction of global motion using Gaussian distributions of directions.

To anticipate, our results show that direction discrimination of the global motion percept is influenced by both the bandwidth of the controlling direction distribution and duration of the stimuli, but not by the paths travelled by individual dots over time. As will be shown later in the discussion, our data are consistent with a line-element model described previously by Williams *et al*. (1984).

METHODS

Stimuli

Stimuli were 256 computer-generated dots plotted on a cathode ray tube (CRT) display with a relatively fast, P4, phosphor. A mask, with a circular aperture 8° in diameter, covered the face of the CRT. This aperture allowed only about 130 of the 256 dots to be visible at any one time. The density of dots was 2.56 dots per square degree of visual angle. Each dot subtended 6'. Luminance of a single dot was about 0.82 cd/m^2. The luminance of the mask was 0.07 cd/m2; the veiling luminance was 0.03 cd/m^2.

Stimuli were presented at a frame rate of 17.5 Hz. From frame to frame, each dot's movements were controlled by a predefined distribution of directions stored as an array of x- and y-increments. The predefined distribution of directions chosen was

Gaussian.[1] The computer read the increment values for a dot's movements from the array, added the increments to the dot's current position and transmitted the dot's new x- and y-position to the CRT display via digital-to-analog converters. The initial screen location of each dot was randomized for each presentation, rendering the pattern of dots an unreliable clue to direction.

Supported and restrained by a chin-headrest, the seated observer viewed the CRT monocularly from a distance of 57 cm. The non-preferred eye was covered by a translucent patch. The height of the CRT was set so that the center of the aperture was at approximately eye level and observers were required to maintain fixation on a dot located at the center of the aperture. Push-buttons connected to the computer initiated each trial and signalled the observer's responses.

Observers

One of the authors (SW) and four university students served as observers for all experiments. Except for SW, all observers were naive to the purposes of the present experiments and had normal, or corrected-to-normal, visual acuity. Those who required corrective lenses wore them for all experiments.

Procedure

Stimuli were presented in a two-alternative forced-choice procedure. Though the durations of the paired test intervals varied from condition to condition, on any single trial the two were always of equal duration. Interstimulus interval was fixed at 500 msec.

Different distributions of directions governed motion in the two intervals of each trial. One test interval, picked at random, was governed by a distribution whose mean direction was 90 deg (upwards); we'll refer to this stimulus as the standard. Motion in the other test interval was governed by a distribution whose mean was greater than 90 deg (that is, counterclockwise of upwards); we'll refer to this stimulus as the comparison. The observer had to identify the interval in which the global direction of motion was upwards.

A session consisted of six blocks, 48 trials each. A block of trials was characterized by one combination of direction bandwidth and test-interval duration. In order to produce a large range of discrimination performance, from chance to near perfection, six comparison stimuli with different mean directions were used in each block. Trial-wise feedback was provided, with a low tone signalling an incorrect response. Approximately four seconds elapsed between trials. Over any 48-trial block, the standard stimulus appeared equally often in the first and second intervals.

EXPERIMENTS

Experiment I. Bandwidth and Duration

This experiment examined direction discrimination as a function of i)the directions present in the stimulus, and ii)stimulus duration. Four ranges of directions were used, each defined by a different Gaussian distribution of directions. The distributions had standard deviations (SD) of 0.0,[2] 17, 34, and 51 deg. Larger standard deviations, or bandwidths, imply a greater range of directions was simultaneously present in the cinematogram. All standard deviations used produced global motion in the approximate direction of the mean of the distribution.

A pilot study showed that discrimination varied with bandwidth. So, to span the psychometric functions of each bandwidth, sets of comparison stimuli with different means were needed. Table 1 lists the six comparison means associated with each bandwidth. Five durations of presentation, three, six, nine, 12, and 25 frames, were completely crossed with the four bandwidths. For each combination of bandwidth and duration, an observer was tested on a total of 288 trials.

Table 1 about here

Analysis

Responses were aggregated to yield the percent correct for each combination of standard and comparison. The percent correct responses for individual observers were then fit by the Quick (1974) psychometric function, given by

$$\Psi(S)=1-2^{-(k*S)^{P}} \qquad [1]$$

where S is the separation in mean direction between the standard and comparison stimulus, measured in deg, 1/k is the difference between standard and comparison means at which $\Psi(S)$ equals 0.5 (chance performance), and P determines the maximum slope of the function in the neighborhood of 75% correct. This function provided good fits to the observed data (mean r^2 for 100 data sets was 0.89). Discrimination thresholds, defined as the difference between standard and comparison mean directions sufficient to yield 75%

correct, were evaluated from the fitted psychometric functions. Threshold values were then treated by analysis of variance (ANOVA) including a trend analysis on the two variables.[3]

RESULTS

Discrimination thresholds, averaged over observers, are plotted as a function of bandwidth in Figure 1. As the figure shows, discrimination thresholds for each duration increased with stimulus bandwidth. Generally, discrimination thresholds changed relatively little as stimulus SD was increased from 0.0 to 17 degrees, but changed substantially with further increases. This observation was confirmed with a trend analysis of the data averaged over durations, which yielded significant linear and non-linear components ($F_{1,2}$ = 5520.72 and $F_{2,4}$ = 8.45, both p<0.05). Notice that at the smallest bandwidths, the discrimination thresholds for the four longest durations are indistinguishable. However divergence does occur as bandwidth gets larger. In contrast, the results at the shortest duration, three frames, differ from those of other durations at all bandwidths. This interaction between bandwidth and duration was confirmed by the ANOVA ($F_{12,24}$ = 13.03, p<.05). This implies that as bandwidth grows, it may take longer to perceive the global flow. It is clear however, that regardless of bandwidth, discrimination thresholds obtained with the briefest presentations are consistently higher than those obtained with longer ones.

Figure 1 about here

To more clearly show the effect of duration, we have replotted the data as a function of duration in Figure 2. The figure shows a progressive decrease in discrimination threshold as a function of duration (linear trend $F_{1,2} = 256.74$, $p<0.05$). However, the decrease in threshold with duration also contains non-linear components ($F_{3,6} = 14.72$, $p<0.05$). A larger decrease occurred when duration was increased from three to six frames than when duration was increased from 12 to 25 frames. Moreover, discrimination thresholds for the two smallest bandwidths seemed to reach an asymptotic level between six and 25 frames of duration. In contrast, for the largest bandwidth, each increase in duration produced a further decrease in the discrimination threshold.

Figure 2 about here

Experiment II. Effective Dot Path

In Experiment I, discrimination thresholds increased as bandwidth increased. However, because several aspects of the stimuli covary with bandwidth, that experiment does not allow unequivocal inferences to be made about the cause of the threshold increase. By definition, the number of directions contained in a stimulus increases with bandwidth. So, as bandwidth increases, the path taken by any single dot contains a greater <u>variety</u> of directions. This greater variety might itself have increased the variability of the perceived global direction, thereby impairing global direction discrimination for the stimulus as a whole. We wanted to determine, therefore, how discrimination performance might vary with the number of directions occurring in each dot's path.

To answer this question, we created two stimuli that produced very different individual dot paths but had the same aggregate direction distribution. Both types of stimuli are illustrated in Figure 3. In one, dots took a two-dimensional random walk as described earlier. Because each dot's path was random, within limits imposed by the distribution bandwidth, we'll refer to such a stimulus as the <u>random-path</u> type. Such paths are represented in panel A for two different dots. In the other type of stimulus, a different scheme generated a dot's path. Once a dot had randomly chosen a direction for its first displacement, it continued to move in that same direction for the entire presentation. Because each dot moved along its own characteristic fixed path, we'll refer to such a stimulus as the <u>fixed-path</u> type. Such paths are represented in panel B for two different dots. Note that although the aggregate direction distributions for both stimuli are identical, the variability of their dot paths are very different. In the <u>random-path</u> stimulus, the controlling distribution of directions creates differences between different dots' paths, and also introduces randomness to any single dot's path. In the <u>fixed-path</u> stimulus, the controlling distribution affects only differences between different dots' paths.

Figure 3 about here

The two stimulus types were used to produce three test conditions. In one condition, both presentations within a single trial were fixed-path stimuli (fixed-path condition). In a second condition, both presentations were random-path stimuli (random-path condition). In the third condition, one random-path and one fixed-path stimulus were presented on each trial (combined condition). In this last condition, the two types of motion were completely crossed with respect to which served as the standard or comparison and also their presentation order.

Discrimination performance was measured for six separations between the standard and comparison mean directions: 2, 4, 5, 6, 8, and 10 deg. All stimuli had a Gaussian direction distribution with a standard deviation of 34 deg. Each stimulus was presented for nine frames. This bandwidth and duration were chosen because in previous experiments this combination produced a moderate level of performance. This ensured some latitude for discrimination performance to improve or grow poorer as condition varied from random-path to fixed-path. Observers were the same as those in Experiment I.

RESULTS

The data, averaged over observers and represented as percent correct, are plotted as a function of the difference in mean direction between the standard and comparison stimuli in Figure 4A. The figure shows that all three conditions yielded similar discrimination ($F_{2,8}$ = 1.22, p>.05).

Figure 4 about here

At the duration used in this experiment, nine frames, the two types of motion were different. However, if one looked at the stimuli through a narrow time window, in particular, examining only a single pair of successive frames, the minimum needed to define motion, the two types of stimuli would be indistinguishable. We were concerned, therefore, that this short-term similarity between stimuli might account for the similarity in performance with the two types of motion. This concern would be serious if performance had become asymptotic at a presentation of just two frames. Then, observers would have extracted all the necessary stimulus information before any real differences between stimulus types could have become manifest. But for our experiments this concern is not justified.

Results from Experiment I show that asymptotic performance in Experiment II would certainly have required presentations longer than just two frames. In Figure 4B we have plotted the average of the earlier results for the stimulus with an SD of 34 degrees presented for three frames, the shortest presentation used. The averaged results from the present experiment, for both stimulus types, are also plotted in that figure. Recall that all cinematograms in that earlier experiment were of the type we've labelled "random path". Note that performance with presentations of only three frames in Experiment I was far below that obtained in Experiment II, with nine frames. Therefore, within just two frames, observers in Experiment II had not extracted all the necessary information to determine the direction of motion. So, the identity of random-path and constant-path stimuli over the first two frames of presentation cannot explain the lack of performance difference between the stimuli at nine frames.

The results of Experiment II suggest that individual dot paths over frames are not being used by the visual system in determining the direction of global perceived motion. Rather, perceived global direction seems to depend only upon the <u>distribution</u> of directions of motion present from one frame to the next. That is, the visual system keeps track of the directions created by any one displacement but does not keep track of the successive movements, over frames, of individual dots.

DISCUSSION

As stated earlier, one of the major objectives of this research is to account for our results with a line-element model of direction discrimination. Before discussing the model, it will be useful to relate our results to those in the literature and discuss the implications that these results hold for research in motion perception.

We have found that direction discrimination of random-dot cinematograms depends upon certain stimulus dimensions. First, increasing stimulus bandwidth decreases direction discrimination. Further, increasing stimulus duration results in an improvement in discrimination performance. However, in developing its representation of global direction, the visual system appears to disregard information about individual dot paths over time.

Williams and Sekuler (1984), using stimuli similar to that used here, found that global motion in a single direction was always seen when the range of the uniform direction distribution was less than or equal to 180 deg. Experiment I showed that, although unidirectional global motion was always perceived, as the bandwidth of the direction distribution increased so did the discrimination threshold. The present results suggest that although coherent global flow can be created by any one of a wide range of bandwidths, the precise direction seen may not be as predictable. In other words, the directional bandwidth controls the precision with which the perceived direction matches the mean of the direction distribution.

Experiment I also provided some indication of the integrative power of the visual system in determining direction of motion. Figure 1 showed that direction discrimination did not change significantly when the bandwidth of the stimulus was raised from SD=0.0 to SD=17 deg. This occurred even though the two distributions produced highly distinguishable patterns of movement. The visual system seems to extract and integrate directional information just as easily from stimuli containing many different directions (the stimulus with an SD of 17 deg contained 79 different directions of motion) as it does with only a single direction present.

But bandwidth was not the only variable that influenced discrimination. Stimulus duration also had an impact: as the duration of the stimuli increased, direction discrimination improved. This implies some sort of temporal summation in the process that governs perceived direction of motion. Note that the number of frames needed to reach asymptotic performance is not the same for all bandwidths: as bandwidth decreases, fewer frames are needed to produce asymptotic performance.

Experiment II examined the effect of dot path on discrimination. The results showed that when direction distributions were identical, whether the dots took random walks or followed fixed but different paths, discrimination was unchanged. Previously, Williams and Sekuler (1984) showed that the global percept of motion does not depend on the spatial relationship between local motion vectors over time. Our findings agree with this view: when many vectors of motion are present, the direction of global motion is determined by the distribution of directions rather than by the individual dot paths.

This result also has some methodological, as well as theoretical, implications. Some researchers, utilizing random dot displays, have purposely limited the lifespan of individual dots to restrict the directional information contained within a single dot path (e.g. Mather and Moulden, 1980; Mather and Moulden, 1983). The present result, that individual dot paths do not affect direction discrimination, suggests that this control may not always be necessary. When the stimulus is comprised of many random dots, the visual system does not necessarily utilize information about the consecutive movements of individual dots.

THEORY

A Line-Element Model of Direction Discrimination

As stated earlier, one of our objectives was to account for global direction discrimination with a line-element model. Line-element models have been successful in accounting for several visual discrimination tasks involving dimensions such as wavelength and spatial-frequency (Graham, 1965; Wilson and Gelb, 1984; Wilson and Regan, 1984; Wilson, 1985). A line-element model has also been useful for predicting the conditions under which random dot displays with very different direction distributions would be metameric, that is indistinguishable perceptually despite their considerable physical differences (Williams et al., 1984).

Any line-element model has three defining characteristics. First, it postulates mechanisms whose sensitivity profiles span the stimulus dimension of interest. For any stimulus, the total response of a mechanism is the sum of that mechanism's individual responses to each component of the stimulus. Second, discrimination between two stimuli depends upon the change in a mechanism's response as a result of a change in stimulus components. Finally, the differences in responses to two stimuli are pooled over all mechanisms. This implies that the discriminability of two stimuli is a function of a scalar value (Graham, 1965).

An example of a line-element model is one Williams et al. (1984) used to predict which set of discrete directions of motion would have to be mixed in order to generate a percept that was indistinguishable from one generated by a stimulus containing a broad band of directions of motion. This model comprised a set of direction selective mechanisms, and the response of the model depended only upon the component directions of the stimulus. Based on the success of this line-element model and the demonstration that direction discrimination depends only upon the distribution of directions, it seemed reasonable to attempt to fit the present data with the same model.

In the remainder of the discussion, we will describe the basic structure of the line-element model that we used to account for the present data. Parameters of the model will be estimated using data obtained for stimuli with Gaussian distributions of directions presented for 12 frames. The same parameters will then be used to account data obtained with different presentation durations and predict results for stimuli that had uniform, rather than Gaussian, direction distributions.

Description of the Model

The basic structure and assumptions of the present model are are the same as those used to account for motion metamers (Williams et al., 1984). The present model assumes that the full range of directions (360°) is spanned by a small number of evenly-spaced, bandlimited, directionally-selective mechanisms. All mechanisms have the same Gaussian profile; center-to-center separation between any two adjacent mechanisms is equal to the half-amplitude half-bandwidth of a mechanism.

The sensitivity of the i^{th} mechanism, centered at θ_i, to direction of motion θ is given by

$$S_i(\theta) = \exp^{\{-[(\theta-\theta_i)/h]^2 * \ln 2\}} \qquad [2]$$

where h is the half-amplitude half-bandwidth of the mechanism. The response of the i^{th} mechanism to a distribution of directions, $D(\theta)$, is given by

$$R_i(D) = \sum_{\theta=1}^{360} S_i(\theta) * pr\{D(\theta)\}, \qquad [3]$$

where $S_i(\theta)$ is the i^{th} mechanism's sensitivity to direction θ, and $pr\{D(\theta)\}$ is the proportion of dots in distribution $D(\theta)$ that move in direction θ.

To predict the discriminability of any two distributions, $D(\theta_1)$ and $D(\theta_2)$, with different mean directions, one calculates the difference, for each mechanism, between its responses to the two distributions

$$\Delta R_i = R_i\{D(\theta_1)\} - R_i\{D(\theta_2)\}. \qquad [4]$$

These differences are then pooled for all the individual mechanisms according to a Q^{th} norm rule:

$$\Delta R = \{\sum_{i=1}^{M} |\Delta R_i|^Q\}^{1/Q}, \qquad [5]$$

where M is the number of mechanisms. ΔR represents the total difference between the responses to the two stimuli generated within the visual system. Note that this method of pooling allows for the effects of probability summation (Quick,1974).

The variable Q determines the way response differences, ΔR_i, for each mechanism will be combined. If Q=1, all ΔR_i's are weighted equally and the system would be taking the simple sum of all ΔR_i's. If Q>1, the larger values of ΔR_i are weighted more heavily than smaller values; if Q=infinity, the model acts as a peak detector, taking only the single largest value of ΔR_i into account.

In order to relate the predicted values of ΔR to the data obtained in Experiment I, we used a psychometric function of the form:

$$Y(\Delta R) = 1 - 2^{-(k*\Delta R)^P} \qquad [6]$$

where k is equal to the value of $1/\Delta R$ at $Y(\Delta R)=0.50$ and P is related to the slope of the psychometric function.

The model as described above has four free parameters, two of which we fixed on *a priori* grounds. Previous researchers, Wilson and Gelb (1984), have shown that when Q=2, a line-element model provides good fits to spatial-frequency discrimination data when the stimuli are presented under *sustained* temporal conditions. The temporal modulation of their sustained stimulus was Gaussian with a 1/e time constant of about 250 msec. Following Wilson and Gelb, we decided to use Q=2 in order to fit the data we obtained at a duration of 12 frames, since at this duration, thresholds for the three smallest standard deviations first reached asymptotic levels. This decision left three free parameters, k, P, and M, the number of mechanisms.

We set M=12 in accordance with Williams _et al._ (1984) who found that a model with 12 mechanisms accounted for metameric relations between cinematograms that contained a wide range of directions and cinematograms that contained just a few directions. Having fixed Q and M, we estimated the optimum values for k and P by a least-mean-squares fit to Experiment I data presented for 12 frames. Table 2 shows the chi-square (χ^2) goodness-of-fit values obtained for best-fits to the present data. All χ^2 values are well below the critical value suggesting that the model fit the data well.

Table 2 about here

Model Fits for Various Durations

The model as described above, provided a satisfactory account of data obtained for stimuli presented for a long duration, 12 frames, with Q=2. Since the six-, nine-, 12-, and 25-frame conditions seemed to be grouped together (see Figure 1), the same parameters used to fit the 12-frame data were also used to fit the six-, nine-, and 25-frame data. The predicted values along with the observed data for the six-frame condition, for all observers, are presented in Figure 5. Those for the nine-frame condition appear in

Figure 6 while those for the 25-frame condition appear in Figure 7. Data are shown by the filled squares and the model by the lines. For all three duration conditions, the model captures the trend of the data. Chi-square goodness-of-fit values for the six-, nine-, and 25-frame data appear in Table 3. The χ^2 values for all observers were below the critical value.

Figure 5, 6, and 7 about here

Table 3 about here

Discrimination thresholds obtained at durations of six frames or greater appear to be grouped together (see Figure 1). However, for the shortest presentation, three frames, discrimination was poorer. Since a model of direction discrimination should account for this effect of duration, we sought to use the present model to predict discrimination for this very short stimulus duration.

Previously, Wilson and Gelb (1984) demonstrated an empirical relation between Q and stimulus duration. They found that a line-element model with Q=2 predicted spatial-frequency discrimination when the stimuli were presented in sustained temporal conditions. When the stimulus was only presented for about 125 msec (transient condition), Q=2 did not give a good account of the data, but Q=6 did. Since the duration of three frames, in msec, was close to that of the transient condition described by Wilson and Gelb, we used Q=6 to predict discrimination in the three-frame condition. The values of k, M, and P remained fixed at the values previously estimated.

Figure 8 compares the model fits to the three-frame data for all observers, measured for various stimulus standard deviations.

Data are represented by filled squares and the model calculations by the lines. Across any row, all the graphs show data for a single standard deviation; within any column, graphs are for a single observer. Table 3 lists the χ^2 values for each observer. Since there were four standard deviations crossed with six separations, there were a total of 24 data points per person used in the calculation of χ^2. As can be seen, all but one of the χ^2 values are below the critical value. Inspection of Figure 8 shows that although the general trend of the data is captured by the model, the fits are not particularly good for the largest standard deviation. The fits would not have been appreciably improved by increasing Q beyond its set value of six since predictions change little as Q is raised above this value. This relatively poor fit to the data can not be reconciled at this time.

Figure 8 about here

Discrimination with Uniform Distributions

We next sought to determine whether the model parameters developed for long-duration stimuli with Gaussian direction distributions (Experiment I) could also account for performance with a different distribution of directions. So we measured direction discrimination, for the same observers as before, now using stimuli with uniform direction distributions. The uniform distributions had ranges of 1, 31, 91, and 161 deg. As we did earlier with the Gaussian stimuli, discrimination was measured for six separations between mean directions, yielding 24 data points per person (separation values for each uniform distribution are found in Table 1).

All stimuli were presented for 12 frames.

Figure 9 compares the predictions of the 12-mechanism model to data obtained with the four uniform stimuli for all observers. This is a parameter free fit to the data, the parameters having been determined in fitting the model to the long-duration Gaussian data. Data are represented by the filled squares and predictions by the lines. Inspection of the figure shows that qualitatively, the model captures the trends in the observed data well. Chi-square goodness-of-fit values were evaluated, for each observer, using all 24 points obtained with the uniform stimuli. The χ^2 values for each observer for the fitted data (Gaussian stimuli) and predicted data (uniform stimuli) are found in Table 2. For all observers, the χ^2 values were well below the critical value. Thus the same parameters that earlier gave a good account of data with long-duration Gaussian stimuli, also give a good account of data with long-duration uniform stimuli.

Figure 9 about here

Summary of Model Results

For all observers, a line-element model with 12 mechanisms and Q=2, provided a good fit to data obtained with Gaussian direction distributions presented for 12 frames. Consistent with the idea that durations of six frames or greater fall into the same group (see Figure 1), the same parameters that provided good fits for the 12-frame data also provided good fits for the six-, nine-, and 25-frame data. For the briefest stimuli, three frames, the model required that Q=6. Finally, the same parameter set estimated for Gaussian direction distributions, presented for 12 frames, did a good job of predicting discrimination with four uniform distributions, presented for 12 frames.

Further Research

This research raises further questions about the ability of the visual system to integrate direction information. Although we have considered discrimination obtained with durations of six frames or greater as a group, it is apparent that for stimuli with large bandwidths there is a systematic change in discrimination with duration (see Figure 1). The present model, though adequate as a first approximation of the integration process, does not account for this bandwidth-duration interaction. Further research is needed to refine the model to account for this effect.

One aspect that has not been touched on here is the integration of information between the two eyes. In the present experiments, all stimuli were presented monocularly. An experiment that could help establish the locus of the integration would be to present part of the distribution of directions to each eye and measure the perceived direction of motion. By varying the relative proportion of the overall distribution shown to each eye and its directional content, we could establish how the visual system integrates motion information between the two eyes and how dissimilar the two stimuli must be before the integration system fails and rivalry results.

Another question of interest is whether color has an effect on the integration of direction information. Recent physiological research has shown that the cells in the Medial Temporal area (MT), which are particularly responsive to complex moving stimuli (Newsome et al. 1986), seem little influenced by color (Livingstone and Hubel, 1987). If MT neurons were involved in the detection and integration of direction information, then one could psychophysically test whether the color of the components of the moving stimuli affect the perceived direction of motion.

A final question concerns the power of the system to integrate various directions. In particular, how similar must component directions of a stimulus be in order for integration to occur? We have shown that people can discriminate the global direction of motion produced by a distribution of directions, with a high degree of accuracy, even when the bandwidth is quite large. However, we also know that if two very different directions of motion are presented simultaneously, the observer perceives both directions of motion but with the separation between them exaggerated (Marshak and Sekuler, 1979). Stimuli similar to ours could be used to examine the continuum between perceiving a single global direction of motion (integration) and simultaneously perceiving several different separate directions of motion (segregation). To explore this continuum, one could present stimuli containing many different directions, sampled at various spacings, and measure whether observers perceived a single global direction.

CONCLUSIONS

To summarize the findings and implications of the present studies: Increasing stimulus bandwidth decreases direction discrimination. Increasing stimulus duration results in an improvement in discrimination performance. In developing its representation of global direction, the visual system appears to disregard information about individual dot paths. A line-element model with 12 mechanisms accounts for direction discrimination for a wide variety of stimulus bandwidths and durations. The model required a systematic change in Q, the parameter that reflects the mode of pooling across mechanisms, to account for the change in discrimination with duration. A Q of 6 was required for the shortest duration while a Q of 2 was required for longer durations. A possible mechanistic way to interpret the change in Q with duration is that as duration decreases, fewer of the mechanisms' responses enter into the pooled, overall response.

Acknowledgment-- This research was supported by a grant from the U.S. Air Force Office of Scientific Research, AFOSR 85-0370.

REFERENCES

Adelson, E. H. and Movshon, J. A. (1982) Phenomenal coherence of moving gratings, Nature, 300, 523-525.

Graham, C. H. (1965) Color: data and theories. In Vision and Visual Perception (C.H. Graham, et al. ed.), pp.414-451, Wiley, New York.

Livingstone, M. and Hubel, D. (1987) Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. Journal of Neuroscience 7, 3416-3468.

Mather, G. and Moulden, B. (1980) A simultaneous shift in apparent direction: further evidence for a "distribution shift" model of direction coding. Quarterly Journal of Experimental Psychology 32, 325-333.

Marshak, W. and Sekuler, R. (1979) Mutual repulsion between moving visual targets. Science 205, 1399-1401.

Mather, G. and Moulden, B. (1983) Thresholds for movement direction: two directions are less detectable than one. Quarterly Journal of Experimental Psychology 35, 513-518.

Newsome, W., Mikami, A., and Wurtz, R. (1986) Motion selectivity in macaque visual cortex. III. psychophysics and physiology of apparent motion. Journal of Neurophysiology 55, 1340-1351.

Quick, R. F. (1974) A vector-magnitude model for contrast detection. Kybernetik 16, 65-67.

Williams, D. W. and Sekuler, R. (1984) Coherent global motion percepts from stochastic local motions. Vision Research 24, 55-62.

Williams, D. W., Tweten, S., and Sekuler, R. (1984) Using metamers to explore motion perception. Supplement to Investigative Ophthalmology and Visual Science 25, 14.

Wilson, H. R. (1985) Discrimination of contour curvature: data and theory. Journal of the Optical Society of America A 2, 1191-1198.

Wilson, H. R. and Gelb, D. J. (1984) Modified line-element theory for spatial-frequency and width discrimination. Journal of the Optical Society of America A 1, 124-131.

Wilson, H. R. and Regan, D. (1984) Spatial-frequency adaptation and grating discrimination: predictions of a line-element model. Journal of the Optical Society of America A 1, 1091-1096.

FIGURE CAPTIONS

Fig.1 Discrimination thresholds (see text for definition) for five durations, averaged over observers, plotted as a function of stimulus distribution standard deviation (SD). Notice that at all SDs, the three-frame thresholds are higher than all others. At the two smallest stimulus SDs, thresholds are identical for durations of six frames or more. For these same durations, thresholds diverge at larger SDs. At the two largest SDs, there seems to be a systematic change in thresholds with duration; thresholds decrease as duration increases.

Fig.2 Discrimination thresholds (see text for definition) for four stimulus distribution standard deviations, averaged over observers, plotted as a function of duration. Note that for the two smallest distribution SDs (filled and unfilled squares), thresholds have reached an asymptotic minimum after a duration of only six frames.

Fig.3 Two types of individual dot motion, random-path (A) and fixed-path (B). Note that only two directions of local motion are present in both A and B and that the vector-sum of the directions is the same in both cases.

Fig.4 Percent correct judgments as a function of mean direction separation. Data are averaged over all observers. (A) Data are presented for three dot-path conditions. Average standard error bars are provided in the legend for each condition. Notice that the three different conditions yield quite similar results. (B) Data, averaged over the three dot-path conditions, are presented with data from Experiment I. These Experiment I data were obtained using the same stimulus bandwidth but presented for only three frames. Standard error bars are provided on each curve. Note that the three-frame data from Experiment I are far below the averaged data from Experiment II .

Fig.5 Data for four stimuli with Gaussian distributions of directions with different standard deviations (SD), presented for a duration of six frames. Data are represented by the filled squares while the solid curves represent fits from a 12-mechanism line-element model with Q=2. Each row of graphs represents data for a single stimulus distribution SD; each column provides a single observer's data. Note that the slope of the data gets shallcwer as the distribution SD increases and that the model fits follow this trend of the data.

Fig.6 As in Figure 5, but for a duration of nine frames.

Fig.7 As in Figure 5, but for a duration of 25 frames.

Fig.8 Data for four stimuli with Gaussian distributions of directions with different standard deviations (SD), presented for a duration of three frames. Data are represented by the filled squares while the solid curves represent fits from a 12-mechanism line-element model with Q=6.

Fig.9 Data for four bandwidths of uniform stimuli presented for 12 frames. Data are represented by the filled squares while the solid curves represent predictions from a 12-mechanism line-element model with Q=2. Model parameters were evaluated from fitting data obtained for four stimuli with different Gaussian distribution standard deviations presented for 12 frames. Each row of graphs represents data for a single stimulus bandwidth; each column provides a single observer's data. As in the previous figures, the slope of the data gets shallower as the bandwidth increases; this trend is captured well by the model predictions.

Footnotes

1. Because of the discrete nature of the display, it was not possible to present a continuum of directions. We approximated a Gaussian distribution by sampling a one degree intervals.

2. The Gaussian distribution with a standard deviation of 0.0 deg signifies motion in which all dots moved in parallel paths in the same direction.

3. The evaluation of discrimination thresholds produced two extremely large values that were substantially different from the others. These extreme values were due to a lack of monotonicity in two observers' data for a particular bandwidth-duration combination. These two values were excluded from the ANOVA conducted on the bandwidth and duration data.

Table 1. Bandwidths and mean directions of stimuli with Gaussian and uniform direction distributions.

Standard Deviations of Gaussian Distributions	Mean Directions: Standard	Mean Directions: Comparison
0.0 deg (unitary motion)	90 deg (upwards)	91, 92, 93, 94, 95, 96
17 deg	90 deg	91, 92, 94, 95, 96, 98
34 deg	90 deg	92, 94, 95, 96, 98, 100
51 deg	90 deg	92, 95, 97, 99, 102, 105

Ranges of Uniform Distributions	Mean Directions: Standard	Mean Directions: Comparison
1 deg (unitary motion)	90 deg (upwards)	91, 92, 93, 94, 95, 96
31 deg	90 deg	91, 92, 93, 94, 95, 96
91 deg	90 deg	91, 92, 93, 96, 99, 102
161 deg	90 deg	92, 94, 95, 100, 105, 110

Table 2. Chi-square values of model fits to four Gaussian stimuli and predictions for four uniform stimuli presented for 12 frames.

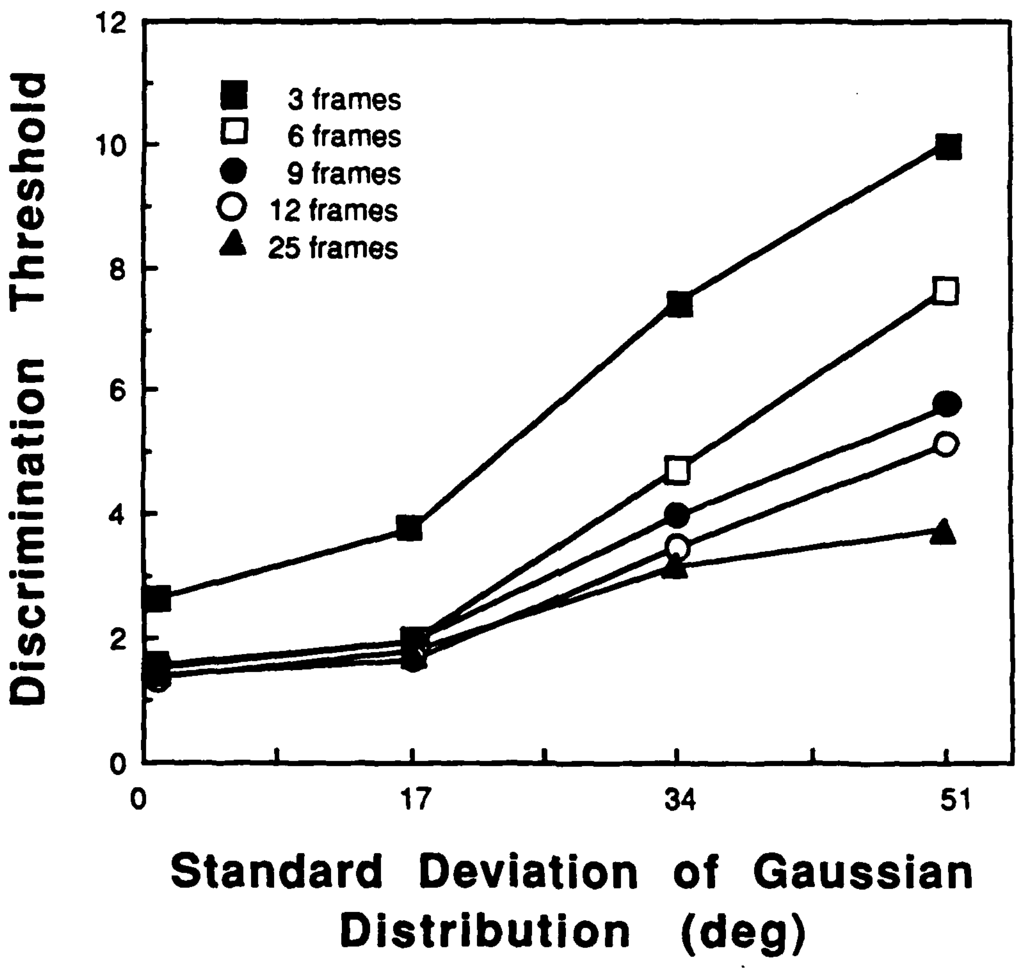
Observer	Gaussian Distribution	Uniform Distribution
CC	8.35	24.92
CP	7.73	9.21
DA	4.88	7.63
JW	12.93	17.38
SW	4.16	5.60
critical $\chi^2_{.95}$	33.9 (df=22)	36.4 (df=24)

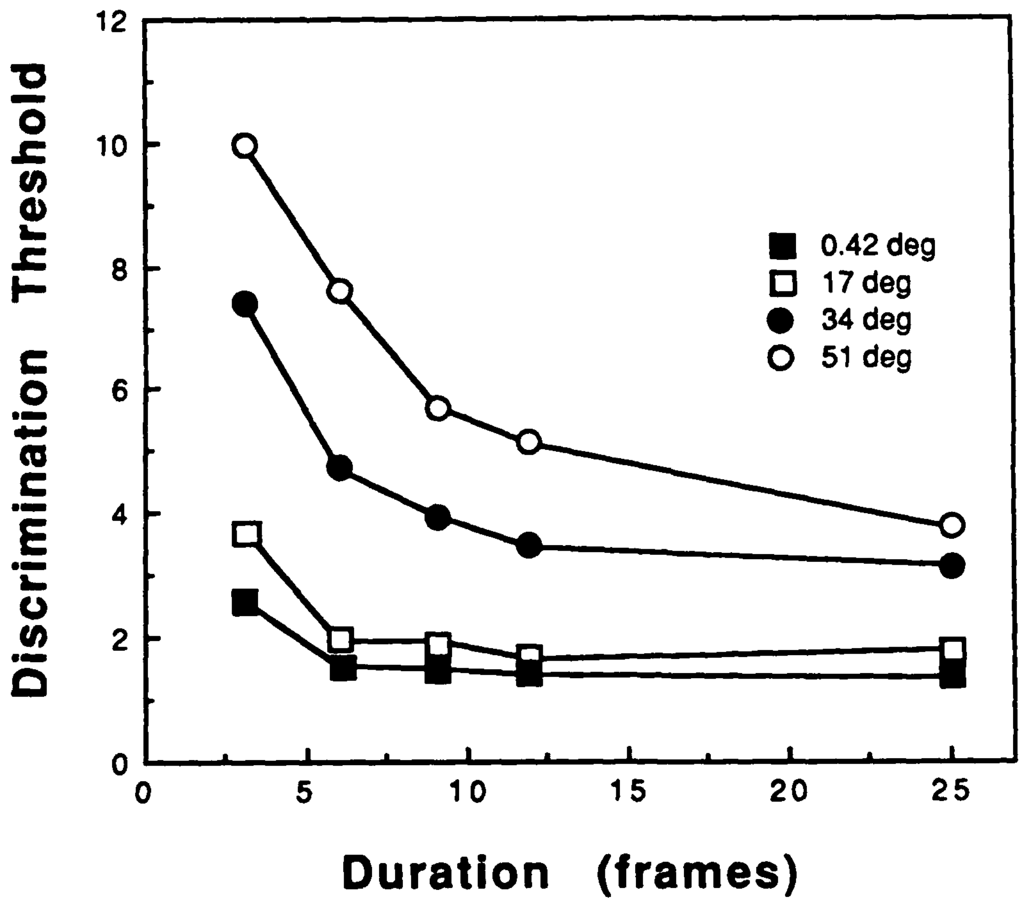
Note: Values underlined exceed critical χ^2.

Table 3. Chi-square values of model fits to four Gaussian stimuli presented for durations of three, six, nine, and 25 frames.

Observer	3 frames	6 frames	9 frames	25 frames
CC	27.62	12.93	11.35	13.03
CP	18.47	10.72	5.15	10.65
DA	24.00	14.66	7.03	5.78
JW	<u>50.17</u>	19.10	7.78	23.40
SW	18.47	15.64	8.35	3.87
critical $\chi^2_{.95}$	35.2 (df=23)	36.4 (df=24)	36.4 (df=24)	36.4 (df=24)

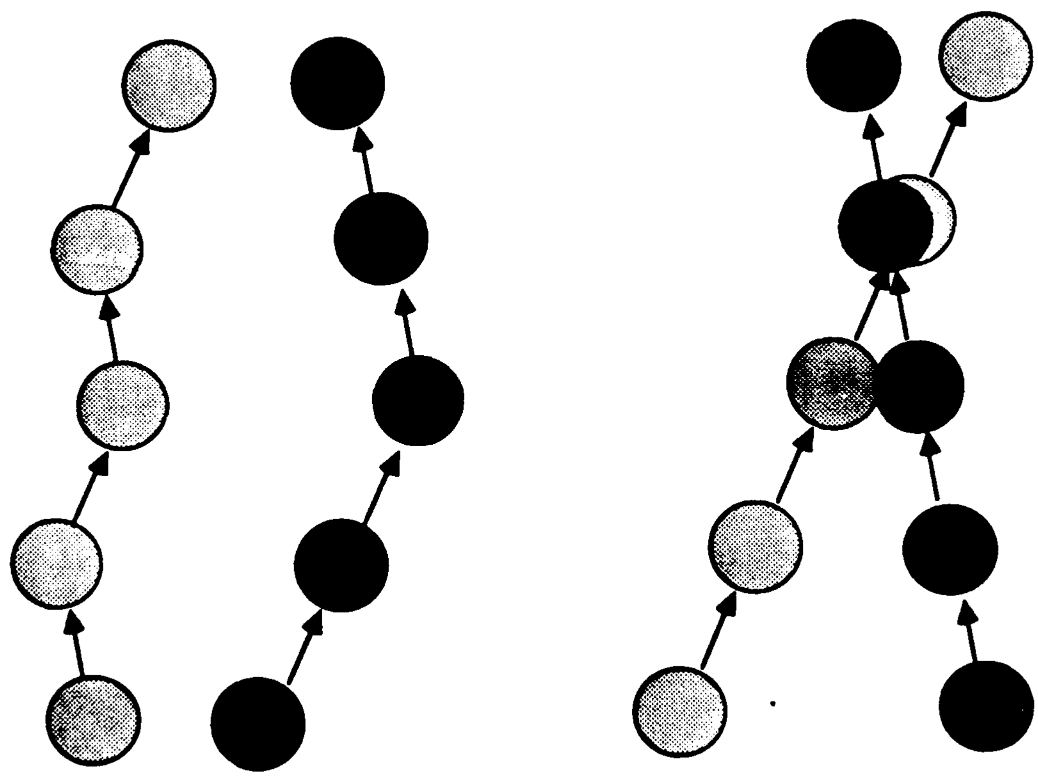
Note: Values <u>underlined</u> exceed critical χ^2.


12
10
8
6
4
2
0
Discrimination Threshold
3 frames
6 frames
9 frames
12 frames
25 frames
0
17
34
51
Standard Deviation of Gaussian Distribution (deg)

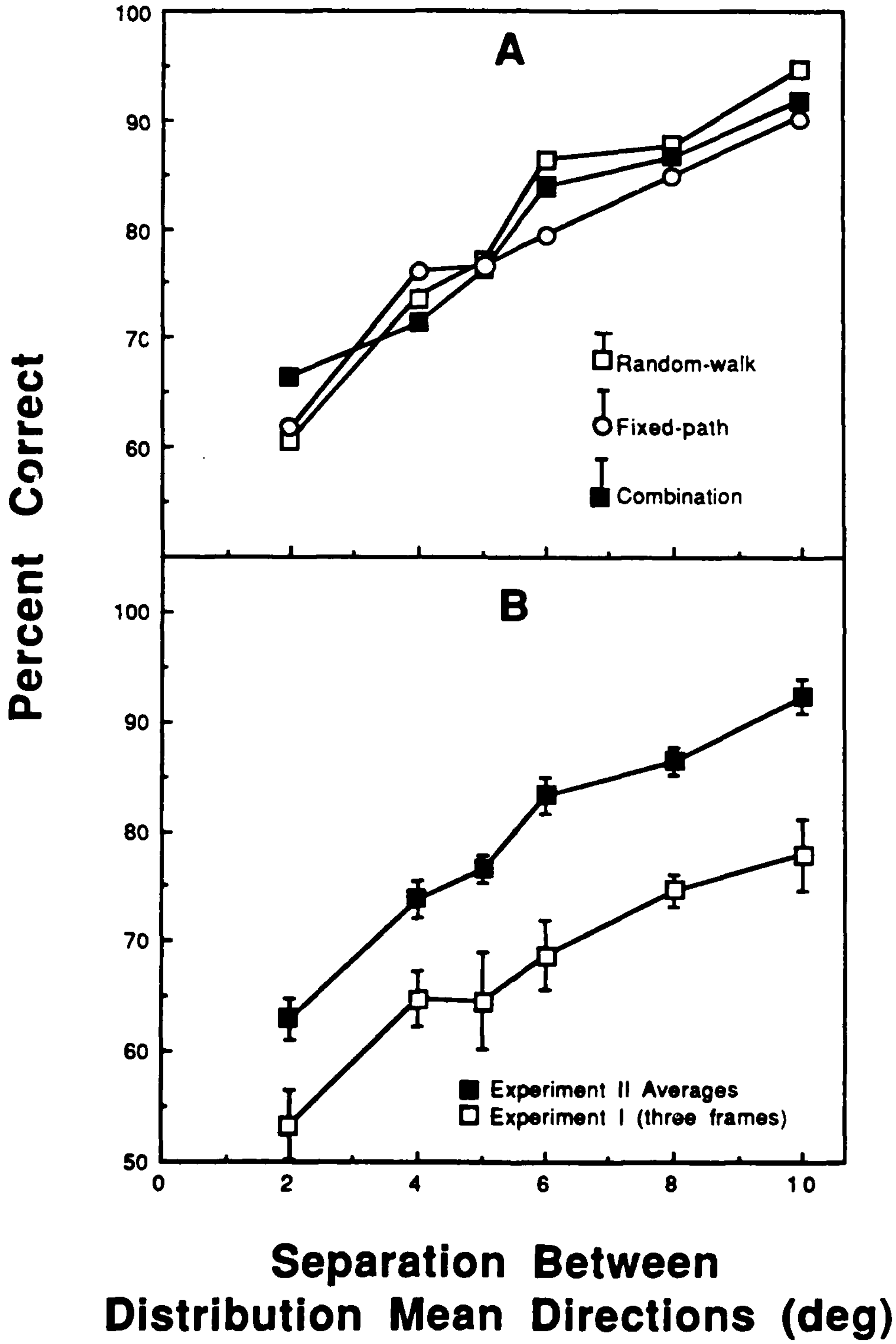


12
10
8
6
4
2
0
0
5
10
15
20
25
Discrimination Threshold
Duration (frames)
0.42 deg
17 deg
34 deg
51 deg

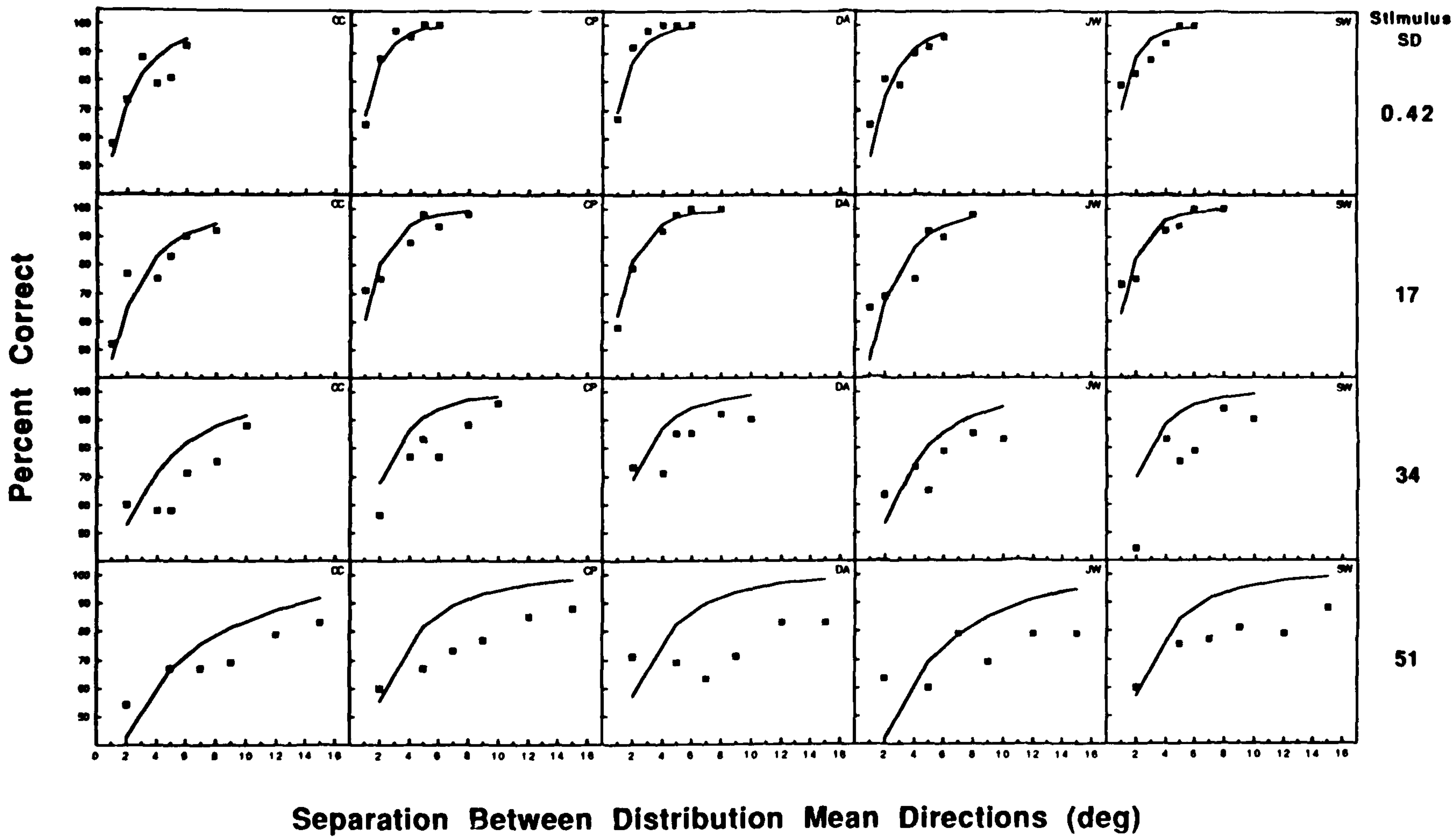
A B



Random-Path Fixed-Path

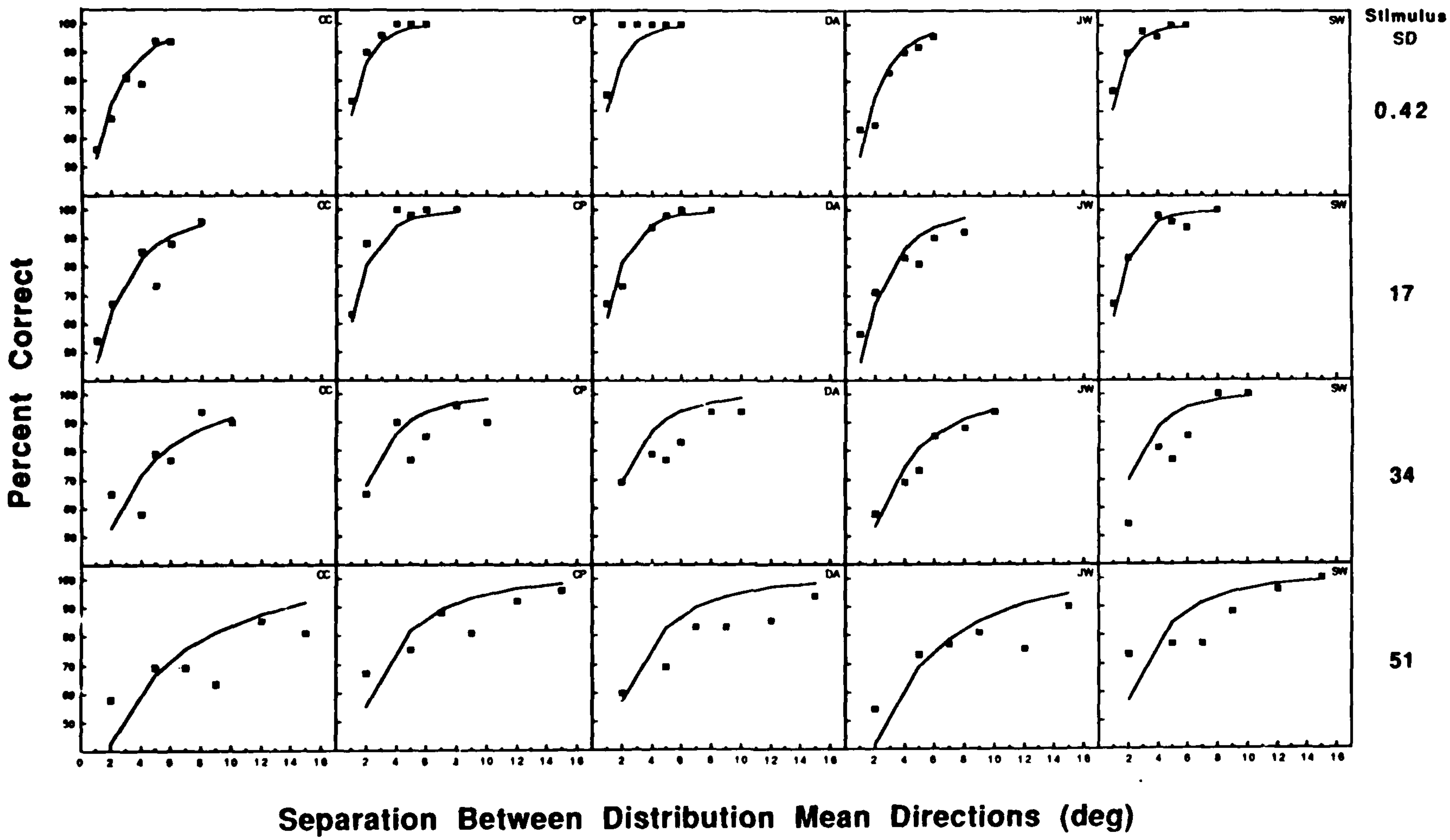


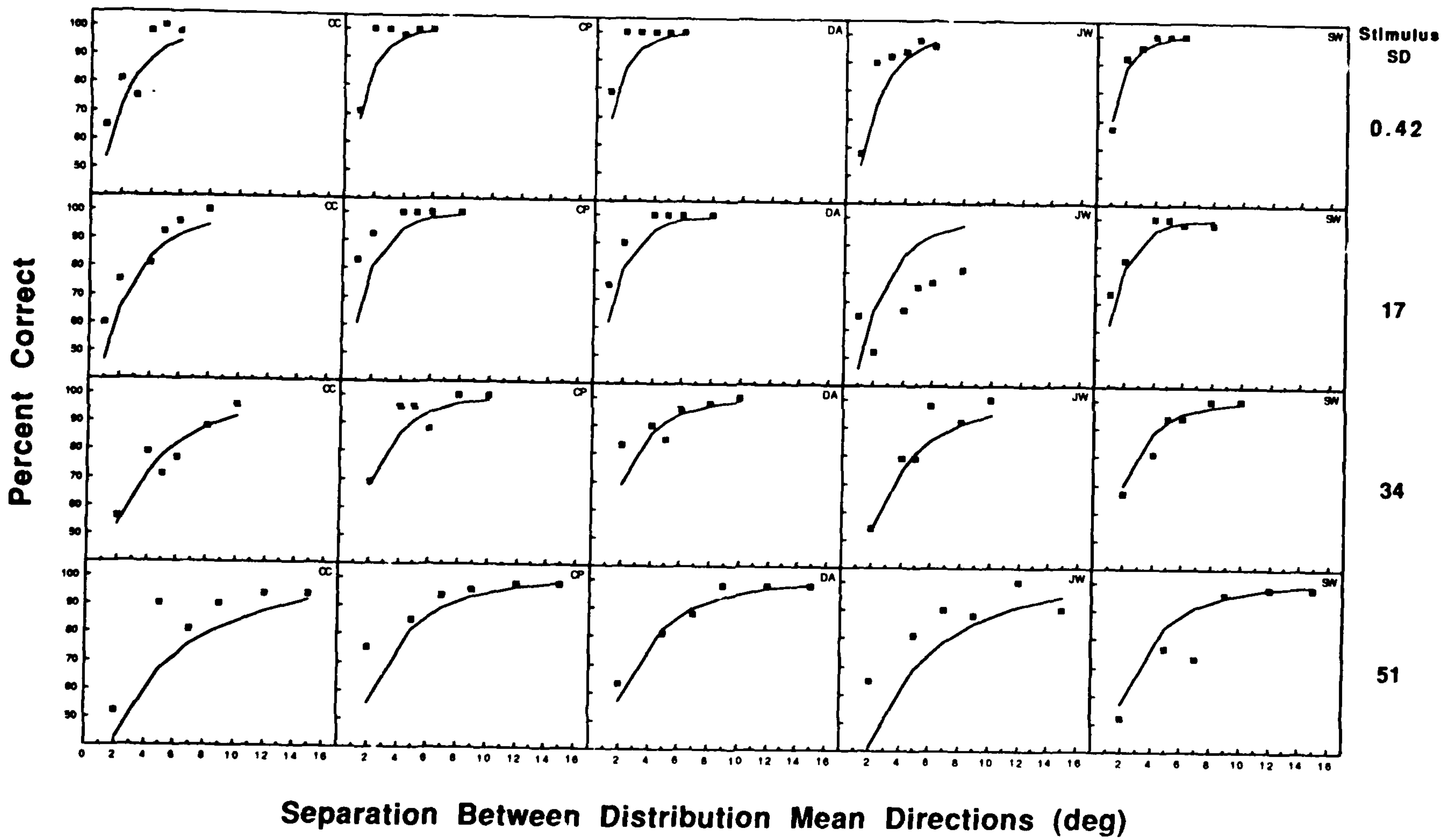
A
B
Random-walk
Fixed-path
Combination
Experiment II Averages
Experiment I (three frames)
Percent Correct
100
90
80
70
60
50
0
2
4
6
8
10
Separation Between
Distribution Mean Directions (deg)

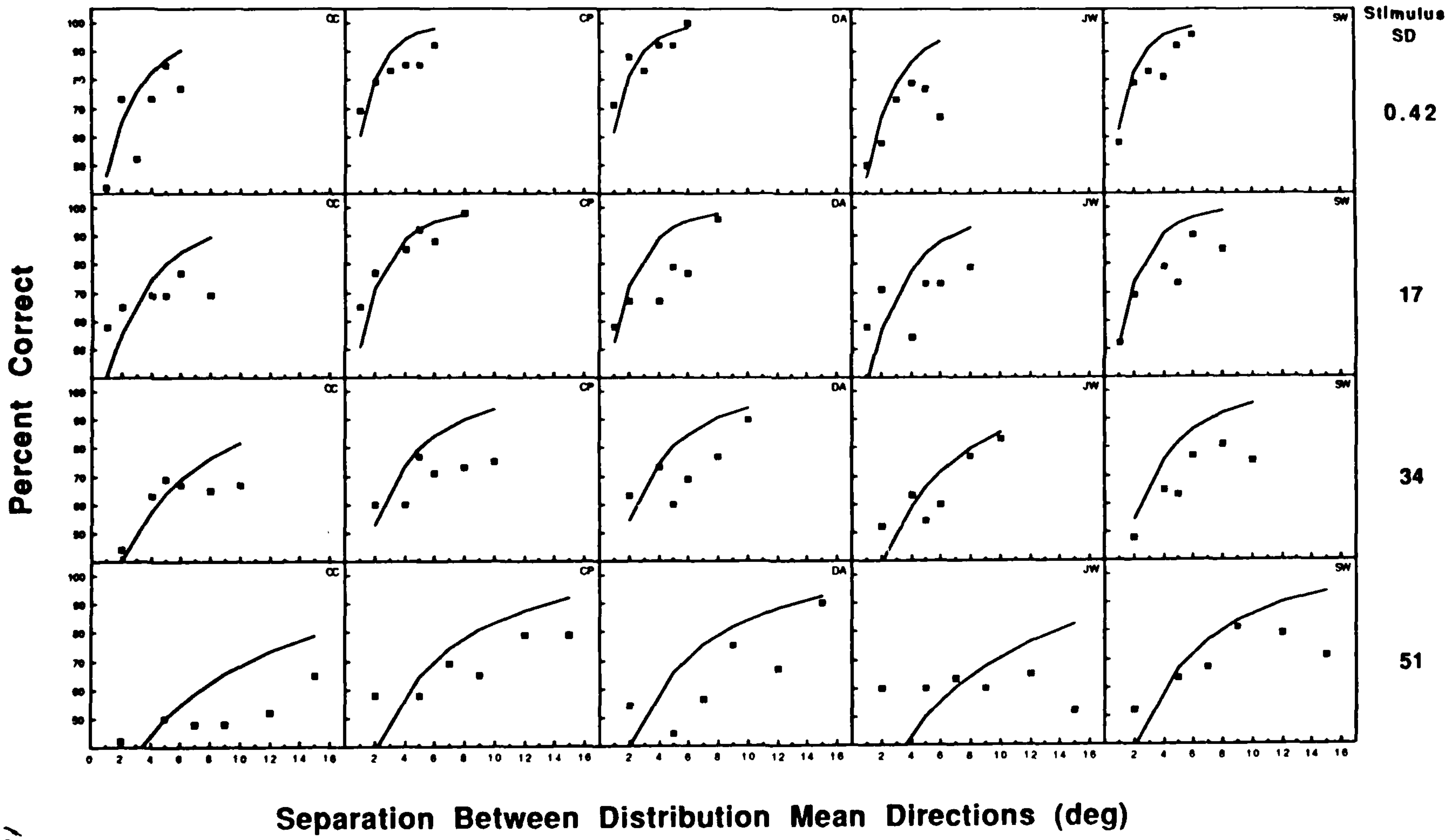


Percent Correct
Separation Between Distribution Mean Directions (deg)
Stimulus SD
0.42
17
34
51
CC
CP
DA
JW
SW

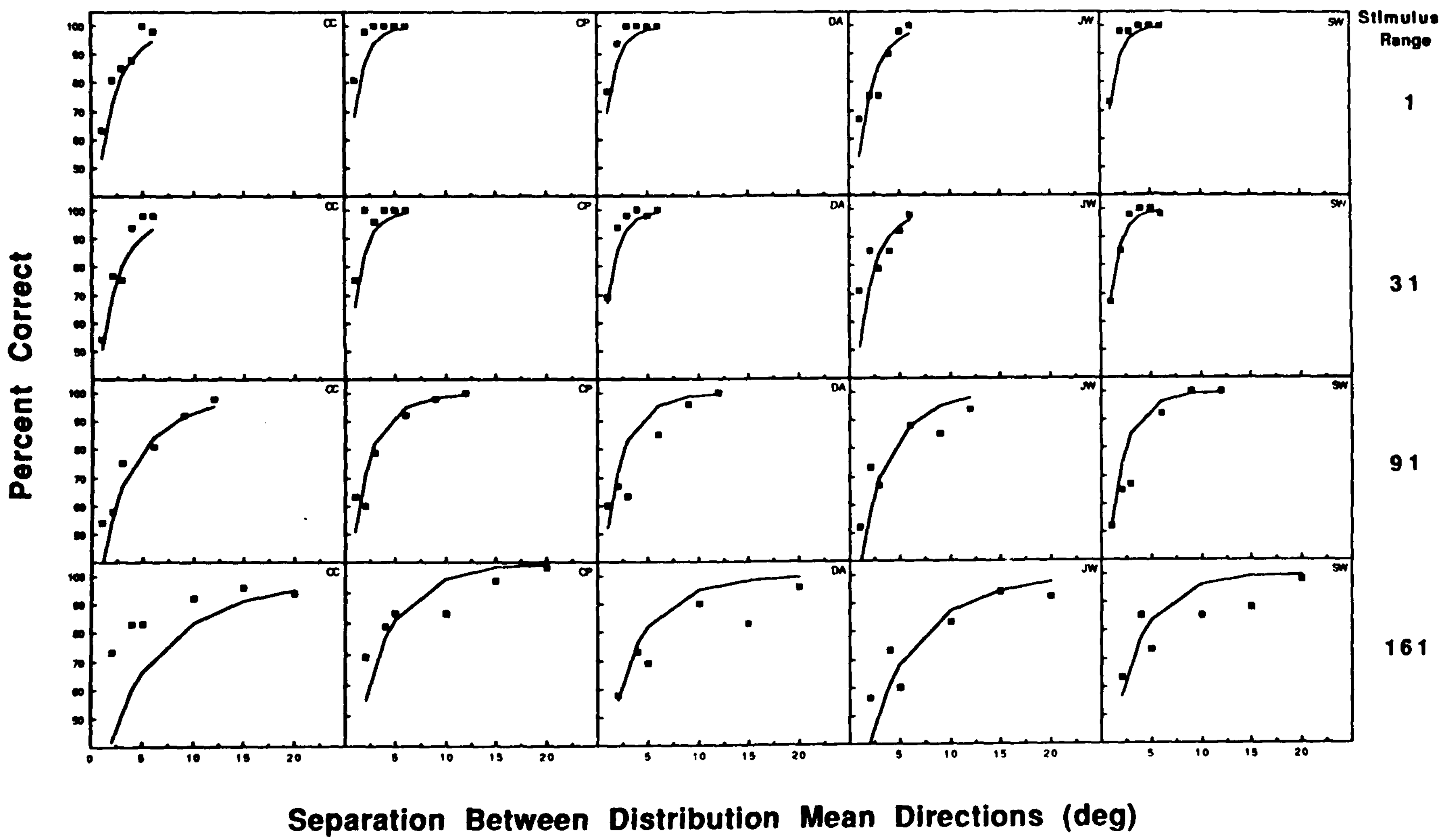
97


Percent Correct
Separation Between Distribution Mean Directions (deg)
Stimulus SD
0.42
17
34
51
CC
CP
DA
JW
SW


Percent Correct
Separation Between Distribution Mean Directions (deg)
Stimulus SD
0.42
17
34
51
CC
CP
DA
JW
SW



Stimulus SD
0.42
17
34
51
CC
CP
DA
JW
SW
Percent Correct
Separation Between Distribution Mean Directions (deg)


Percent Correct
Stimulus Range
1
31
91
161
CC
CP
DA
JW
SW
Separation Between Distribution Mean Directions (deg)

Project Four:

Reaction times to change in speed and direction of motion

Robert Sekuler, Ehtibar Dzhafarov, and Juri Allik

INTRODUCTION

Dzhafarov and Allik proposed the Local Dispersion Model (LD-model) as a framework for interpreting detectability of planar rigid motion with an arbitrary time-position function (Dzhafarov et al., 1981; Dzhafarov, 1982; Dzhafarov et al., 1983; Dzhafarov and Allik, 1984). Predictions from the LD-model were consistent with data on kinematic thresholds and psychometric functions. Of particular importance for the present work, Allik and Dzhafarov (1984) found good quantitative agreement between their model and reaction times (RTs) to motion onset. Consistent with empirical findings (Ball and Sekuler, 1980; Tynan and Sekuler, 1982), the model predicted longer RTs to onset of slow motion than to fast motion.

In those studies of RT to motion onset, after some rest period the stimulus started to move with constant velocity . Now we have measured RTs in a more general situation: a target moves at a constant velocity for some random time, after which its velocity abruptly changes to another constant value. Observers must react as soon as the change in velocity is noticed. Our aim was to develop a theory that would account for the dependence of RT on the relationship between the two velocities. Figure 1 shows the various types of kinematic functions we used. The two phases of motion always had the same, horizontal, orientation; either they differed in speed (panels a and b), or they were in opposite directions (panel c). For each pair of velocities we analyzed mean RTs and standard deviations of RTs. Note, in Figure 1, that the first velocity of a pair sometimes took a zero value (panel a1); in such a case the *change of velocity* is identical to the *onset of uniform motion*, the condition used by Tynan and Sekuler (1982).

[Insert Figure 1 about here]

Dzhafarov and Allik originally designed the LD-model to explain how the visual system distinguishes between a target's motion and non-motion. The model did not deal with detection of change in a particular parameter of motion, *e.g.* a change in direction or a change in speed. However, as we will show in this paper, a simple modification enables the LD-model to predict detectability of changes in velocity.

We will also show that one alternative model for RT to motion onset (Ball and Sekuler, 1980; Tynan and Sekuler, 1982; Allik & Dzhafarov, 1984) fails in the general case of velocity change. This alternative model asserts that reactions to motion onset are initiated when the target has moved through some constant, or critical, distance. The model is therefore referred to as a Constant Distance Model (CD-model).

In testing the models -- Local Dispersion and Constant Distance types -- we were primarily interested in quantitative predictions, and in the plausibility of their parameters' optimal values. Since there is theoretical interest in the way vision encodes direction and speed (for review: Sekuler, 1975; Nakayama, 1985), we also wanted to know whether a single framework could handle RTs to direction reversals as well as RTs to unidirectional speed changes.

Before turning to the details of our empirical research and theoretical analysis, consider a general postulate common to all theoretical treatments of RTs. The postulate is that reaction times are comprised of two additive components. One component, the decision time (t_D) is a function of stimulus parameters such as velocity; the other component is residual time (t_R), the minimum time an observer needs to execute the required response. So, all models considered here agree that

$$RT = t_D + t_R \qquad [1]$$

The various models differ only in their interpretations of the t_D component.

The rest of the paper is organized as follows. First, we briefly discuss the LD-model and the CD-model as formulated for motion detection and for RTs to motion onset. There are two reasons for this discussion. First, these models are prototypes that we are going to transfer to the domain of velocity change; second, the onset of uniform motion is a particular case of velocity change, namely when the first of the two velocities is zero. We shall see that this subset of data forms a strong basis for evaluating the models. After the discussion of the original models, we present some plausible modifications for the situation investigated in our experiments. All the models will be formulated in strictly psychophysical terms: the characteristics of motion on which the decision is based, and the decision rule itself. After the models have been presented, experimental results will be described, and confronted by the models. Finally, the Discussion section considers one biologically plausible system of mechanisms able to extract the required characteristics from the stimulus.

LD-MODEL, CD-MODEL, AND PROPOSITION OF IDENTITY

The Local Dispersion Model (LD-model) has been described in more detail elsewhere (Dzhafarov, 1982; Dzhafarov and Allik, 1984; Dzhafarov *et* *al*.,1983). Consider a two-dimensional luminance profile, $L(x,y)$, whose position changes over time according to some arbitrary kinematic function, $k(t)=\langle k_x(t),k_y(t)\rangle$. The LD-model identifies two separable factors that limit motion detectability. One factor is spatio-temporal luminance fusion (or smearing) along the trajectory of motion; the other factor is a particular characteristic of the kinematic function, its "local dispersion".

Luminance fusion can take place if the kinematic function, $k(t)$, is a high-frequency oscillation, and/or if the moving profile, $L(x,y)$, has a repetitive structure. In either case we have high-frequency luminance flicker at every point of the motion trajectory. Adjacent flickers can fuse in a non-independent fashion because of spatio-temporal luminance integration in the visual system and in the display device. Whether the complete fusion takes place depends on both the kinematic function and the moving profile. If fusion is only partial, or it does not occur at all (as with the leading edge of a unidirectionally moving contour), then detectability of motion depends on the kinematic function only.

The model asserts that the detectability value is given by a moving average over the moving variance of the kinematic function, a value termed Local Dispersion (LD):

$$LD(t) = 1/(2T\tau^2) \int_{t-T}^{t} \int_{t_0-\tau}^{t_0} \int_{t_0-\tau}^{t_0} E[k(t_1),k(t_2)]^2 dt_2 dt_1 dt_0 \qquad [2]$$

where E is the Euclidean distance, τ is the time span of the moving variance (over the stimulus' kinematic function), T is the time span of the moving average (over the moving variance). Note that the term "local" in the name of the model has a temporal rather than a spatial meaning: the LD-value is defined at every moment of time.

Equation 2 means simply that motion detectability is proportional to an average dispersion, or scatter, of a target's temporally close spatial positions. The local dispersion reflects the variance of spatial positions measured within a travelling temporal window, $[t_0 - \tau, t_0]$, and assigned to every moment t_0. At any moment, t, the LD-value is the mean of the moving variance between times t and t-T. Thus if T is zero, motion detectability is proportional to the maximal value of moving variance; if T is infinitely large, detectability depends on the grand mean of all variance values. Zero and infinity form the poles between which the actual value of T lies. Empirically, the ratio T/τ has been found to be a constant, 2, for all the data known to be relevant, though τ does vary with the display conditions and from one observer to the next. The τ is close to 0.5 sec for foveal absolute motion (*i.e.* one without a stationary reference near the motion). Figure 2 illustrates one of the computational algorithms that are equivalent to equation [2]. It will be discussed in more detail in Discussion.

[Insert Figure 2 about here]

Equation [2] represents LD as a particular characteristic, or feature of the stimulus' kinematic function; as a result it has the same ontological status as speed, distance, or acceleration. However the definition of a stimulus parameter on which the subjects might base their choice between "motion" and "no motion", constitutes only the first part of a complete psychophysical model. In the second part one should specify the decision rule for the particular experimental task. Thus, for experiments with kinematic thresholds, like minimum amplitudes of oscillatory motions, one should assume that the motion is detected when the LD-value exceeds some critical level, C^2, where C is a distance-dimensioned parameter (notice that the LD is measured in squared distance units, *e.g.* min^2).

In using the LD-model to predict RTs one needs an assumption that links values of LD to the actual initiation of a reaction. Here again the simplest assumption is that a decision to react is made as soon as LD exceeds some critical value. In applying the LD-model to reaction times elicited by onset of motion, Allik and Dzhafarov (1984) showed that decision time, t_D, can be found from the equation:

$$V^2 t_D^4 (1-3t_D/5\tau)/(12T\tau) = C^2 \qquad [3]$$

V is the motion velocity; T, τ, and C have the same meaning as above. The t_D in equation [3] can be shown to be a decreasing function of V.

For RT experiments the LD-model gave numeric values of T, τ, and C that were similar to the values needed to account for kinematic thresholds and psychometric functions. This similarity is important. It means that in a reaction time experiment an observer actually obeys the experimenter's instructions, initiating reaction as soon as motion is *detected*. Putting it in other words, the similarity of parameters across experimental situations implies that an observer in a reaction time experiment uses the same criterion that an observer would use when kinematic thresholds were being measured. This implication, which we call the *Proposition of Identity*, suggests that reaction time experiments should be considered as a one class of motion detectability experiment. Although they deal with motions well above threshold they reveal the same processes as do other types of experiments on motion detectability.

Only one other model has been applied to data on RT to motion onset, the Constant Distance (CD) Model (Ball and Sekuler, 1980; Tynan and Sekuler, 1982). It states that reactions to motion onset are initiated when the target has moved through some critical distance. When the motion has a constant velocity, V, the decision time, t_D, can be found from the simple formula

$$t_D = \Delta / V \qquad [4]$$

where Δ denotes the critical distance.

It's hard to formulate the Proposition of Identity for the CD-model because the model itself fails with data on kinematic thresholds. Except for oscillatory motion in a middle-frequency range (1-7 Hz), amplitude thresholds are not constant, and even over this limited range the "constant" varies with type of oscillation (Dzhafarov *et al.*, 1981). Nevertheless, some authors insist that the constant displacement rule does hold for very brief unidirectional motions (Cohen and Bonnet, 1972; Johnson and Leibowitz, 1976; Bonnet, 1977, 1982). If this suggestion were even approximately true, then the greatest precision in estimating critical displacement would be reached in the briefest possible motion, namely, an *instantaneous shift of position*. Then the Proposition of Identity for the CD model would reduce to the assumption that the parameter Δ, in equation [4] for reaction time, is close to the threshold for position shift.

THREE MODELS FOR RT TO VELOCITY CHANGE

We have described how the LD-model and the CD-model can account for RTs to motion onset. When uniform motion follows a rest period the solution is given by equations [3] and [4], in combination with the assumption expressed by equation [1]. We will now consider how these formulations can be modified for the general case of change from one velocity, V_0, to another, V_1. Recall that V_0 is the velocity of the first phase of motion that lasts for some random period and then abruptly changes to the second phase, with velocity V_1. The two motion phases have the same orientation, but different absolute values (speed) or signs (direction). Formally speaking, we seek to express RT as a function of $\langle V_0,V_1\rangle$. From the original models we know a part of this function, the dependence of RT on pairs of the type $\langle 0,V\rangle$.

One simple solution suggests itself: reduce the general problem to the particular case for which the solution is already known. Specifically, assume that detection of velocity change, $\langle V_0,V_1\rangle$, is structurally equivalent to detection of onset in the derived motion, $\langle 0,V_1-V_0\rangle$. By *structurally equivalent* we mean identical except for the values of the models' free parameters. Applying this scheme to both CD-model and LD-model, we get generalizations of equations [3] and [4].

For the Constant Distance Model:

$$t_D = \Delta(V_0)/|V_1-V_0| \qquad [5]$$

For the Local Dispersion model:

$$|V_1-V_0|^2 t_D^4(1-3t_D/5\tau)/(12T\tau) = C(V_0)^2 \qquad [6]$$

For both equations, decision time depends upon an *equivalent* velocity rather than upon a directly measurable one. Therefore, as a reminder, we'll label the resulting models with the term "equivalent." So equation [5] describes the *equivalent* Constant Distance Model (eCD-model); equation [6] describes the *equivalent* Local Dispersion Model (eLD-model). The sign of V_0 can be always taken as positive, whereas the sign of V_1 is positive when the two phases are unidirectional, and negative when they have opposite directions.

In both models, Δ and C are functions of V_0, whereas t_R, as usual, is an independent random variable. Although it is not logically necessary, we assume that the parameters T and τ in the LD-model are unmodified by V_0. Moreover we will assume that the values of T and τ are the same as in motion detection experiments. Note that the second assumption is derivable from the first assumption together with the Proposition of Identity.

There is an alternative, perhaps more natural, way to generalize the Local Dispersion Model to the case of velocity change. Provided the first phase of $\langle V_0,V_1\rangle$ lasts long enough ($\approx T+\tau$, estimated as 1.5 sec), LD will stabilize at $LD_0 = V_0^2\tau/12$ (Allik & Dzhafarov, 1984). Then, as velocity changes from V_0 to V_1, the value of LD also will change. We can postulate that velocity change will be detected when the difference between the current LD-value and the initial level LD_0 reaches some critical value. The critical value would depend, in general, on the LD_0 or, equivalently, on V_0:

$$|LD(t_D) - LD_0| = C(V_0)^2 \qquad [7]$$

The difference $LD(t_D) - LD_0$ is given explicitly in the following formula:

$$LD(t_D) - LD_0 =$$
$$(V_1-V_0)t_D^3\{V_0\tau/6 + (V_1-2V_0)t_D/12 - (V_1-V_0)t_D^2/(20\tau)\}/T\tau \qquad [8]$$

Here again, V_0 is taken to be positive, and V_1 is positive if it and V_0 are in the same direction, and negative otherwise. Unlike the alternative version of LD-model discussed earlier (the eLD-model), the local dispersion model in equation [8] can be applied directly to the stimulus' actual, untransformed kinematic function. The only modification in the model is in the decision rule, which is a generalized version of one originally proposed by Allik and Dzhafarov. Therefore we can refer to a *generalized* Local Dispersion, or gLD, model.

EXPERIMENTAL PROCEDURE

The display consisted of 200 spatially-random, bright dots presented under computer control on a large, dim x-y cathode ray tube screen. The dots were 6 min in diameter, and dot-background contrast was set at 4-5 times threshold. The background luminance was about 1.5 cd/m^2. At the start of each trial, the dots appeared and began moving inside a 16 deg diameter circular aperture (see Figure 1). The dots moved horizontally in fixed spatial phase along parallel paths. When a dot reached the edge of the display it wrapped around, reappearing sometime later at the opposite edge. The dots' velocity was controlled by the size of steps, or displacements, from one frame to the next, keeping frame rate constant at 100 Hz. A new set of spatially random dots was generated on each trial.

The experiment consisted of 35 different conditions, each corresponding to one velocity pair, $\langle V_0, V_1 \rangle$. They were tested one at a time in blocks of 50 trials. Over the entire study, each condition was tested on three different occasions, giving in total 150 trials per pair of velocities. The duration of V_0, or stimulus foreperiod, varied according to a uniform random distribution ranging from 1 to 2 seconds. Trials were initiated by the observer.

In thirty conditions, movement during both phases was in a rightward direction. In all these conditions, the subject reacted to a change in *speed* only (Figure 1a,b). Velocity pairs were chosen as pairs from the set of 0 (stationary dots), 1, 2, 4, 8 and 16 deg/sec, with the constraint that the two velocities in a condition could not be the same.

In another five conditions, speeds during both phases were the same. In these conditions, rightward motion during the foreperiod changed abruptly to leftward motion, with no change in speed (Figure 1c). In all these conditions, the subject reacted to a change in *direction* only. Speeds were 1, 2, 4, 8 and 16 deg/sec.

In addition we carried out an auxiliary experiment in order to find out whether any of the obtained results could be specifically associated with our choice of the number of dots in the display, 200. This experiment consisted of 39 different conditions, each corresponding to one of 13 velocity pairs, $\langle V_0, V_1 \rangle$, and one of three dot densities: 50, 100, or 200 dots per screen. A subset of the velocity pairs used in the main experiment was used here: ‹0,1›, ‹0,4›, ‹0,16›; ‹1,8›· ‹2,1›; ‹4,0›, ‹4,16›, ‹4,-4›; ‹8,4›; ‹16,0›, ‹16,1›, ‹16,2›, ‹16,-16›, where the minus sign indicates leftward motion. In all other respects the auxiliary experiment was identical to the main one.

During analysis of the data, all responses less than 100 ms or greater than 1000 ms were discarded, as premature or indicative of the observer's momentary distraction. The number of discarded trials was fairly constant for all conditions, and constituted less than 5% of trials. Remaining trials were used to calculate arithmetic means and standard deviations of RTs for each condition.

One of the observers in the main experiment was an author of this report (RWS); the other observer (JF) was naive with respect to the purposes of the study. A third observer (JLM), also naive, served in the auxiliary experiment.

RESULTS

Figures 3 and 5 show the mean RTs for subjects JF and RWS, respectively. Figures 4 and 6 show corresponding standard deviations of RTs. All panels in every figure contain full set of data, for all $\langle V_1, V_0 \rangle$ pairs, but in each panel the data corresponding to one value of V_0 are "highlighted" (shown by squares). The data are plotted against two abscissae. The lower abscissa represents a measure of similarity between V_0 and V_1, namely $1/|V_1-V_0|^{0.5}$, arrayed linearly. Corresponding values of the difference $|V_1-V_0|$ are shown in the upper abscissa. The square-root operation in our similarity measure has been chosen to linearize the theoretical curves produced by one of the models, as discussed below.

[Insert Figure 3 about here]

[Insert Figure 4 about here]

[Insert Figure 5 about here]

[Insert Figure 6 about here]

One can notice the following main characteristics of the data.

(1) For a fixed V_0, means and standard deviations of the RTs both decrease as the difference between V_1 and V_0 increases.

(2) For a fixed value of $|V_1-V_0|$, RT means and standard deviations increase as the fore-speed, V_0, increases from 4 to 16 deg/sec. With slower forespeeds (between 0 and 4 deg/s) no such trend is discernible.

(3) In ordering both means and standard deviations of RTs, only absolute value of velocity difference, $|V_1-V_0|$, matters, irrespective of whether it represents velocity increment, velocity decrement, or direction reversal. Thus, means and standard deviations of RTs for the velocity pairs $\langle 4,0 \rangle$ and $\langle 4,8 \rangle$ are about the same, and fall between the corresponding RT moments for $\langle 4,16 \rangle$ and $\langle 4,1 \rangle$ (difference in velocities for the first two pairs is 4, for the second 12, and for the third 3 deg/s). In the direction reversal condition $|V_1-V_0|$ is equal to $2V_0$. For example, the difference in velocities for the pair $\langle 16,-16 \rangle$ is equal to 32 deg/s. Therefore, in compliance with the general pattern, the first two moments of the corresponding RT should be less than those for the pair $\langle 16,0 \rangle$.

[Insert Figure 7 about here]

The scattergram in Figure 7 presents the results of the auxiliary experiment in which we varied the number of dots in the display (only mean RTs were analyzed for this experiment). The abscissa represents the mean RTs found with 200 dots in the display for various pairs of $\langle V_0,V_1 \rangle$. Against each mean RT obtained with 200 dots we have plotted the mean RT from the same $\langle V_0,V_1 \rangle$

condition obtained with 50 dots (crosses) and 100 dots (squares). The diagonal line represents the expected loci of data points if mean RT did not differ at all with number of dots in the display. The Friedman rank sums test shows that the difference between the 200 and 100 dot displays, on one hand, and the 50 dot display, on the other is significant ($0.025<p<0.05$). However, it is obvious from the figure that the fourfold change in dot density has a remarkably small effect on mean RT. Therefore our principle results are probably not restricted to the particular number of moving dots used in the main experiment.

Notice that characteristics (1) - (3) of the data are not sufficient to derive ordinal-scale predictions about velocity pairs with different values of both V_0 and $|V_1-V_0|$. A quantitative, model-bound analysis is needed for this purpose; such an analysis follows.

ANALYSIS

COMPUTATIONAL FORMULAS FOR E[RT] AND S[RT]. Formulas [5]-[8] (in combination with formula [1]) do not by themselves allow one to compute RT means and standard deviations. The formulas contain random variables with unknown distributions, t_R and $\Delta(V_0)$s (in the eCD-model) or $C(V_0)$s (in both versions of LD-model). For every combination of these parameters' values one can compute, using the formulas, a single value of RT. What we need instead is a theoretical prediction of RTs' first two moments, expected value, E[RT], and standard deviation, S[RT], for each pair $\langle V_0,V_1\rangle$. Since all models treat RT as a sum of decision time, t_D, and residual time, t_R, the task is reduced to finding the first two moments for the summands, $E[t_R]$, $S[t_R]$, and $S[t_D]$s and $E[t_D]$s, for each pair of velocities, $\langle V_0,V_1\rangle$.

$$E[RT(V_0,V_1)] = E[t_D(V_0,V_1)] + E[t_R]$$

$$S[RT(V_0,V_1)] = \{ S[t_D(V_0,V_1)]^2 + S[t_R]^2 \}^{1/2} \qquad [9]$$

The derivation of expressions for $E[t_D]$ and $S[t_D]$ in the eCD-model is straightforward. However, it's harder to derive exact computational formulas for t_D in the eLD- and gLD-models. These derivations require explicit assumptions about the distribution of parameter C. Because this would add extra free parameters, we wanted to avoid making such assumptions. Instead, we used approximate rather than exact formulas for the eLD- and gLD-model.

The required computational formulas for all the models are given in the Appendix. To account for mean RTs one has to adjust: (1) the value of $E[t_R]$; and (2) a measure of central tendency of distance-dimensioned parameters (Δ or C) corresponding to each value of V_0. To account for standard deviation of RTs one has to adjust: (1) the value of $S[t_R]$; and (2) a measure of variability of distance-dimensioned parameters (Δ or C) for each value of V_0. As we see, the number and the interpretation of the free parameters are identical in the three models. However the measures of central tendency and variability in these models are different. They are shown in Table A1 of the Appendix.

FITTING THE MODELS. There seems to be no conventional statistical procedure to estimate goodness of fit for both means and standard deviations, unless one makes explicit assumptions concerning the distributions of RTs. As explained before, we wanted to avoid assumptions that would add free parameters. Our aim was to determine whether one of the three models provided an account of the data that was substantially better than offered by the other models. This

The values are given in percentage terms in accordance with formula [10]. Thus, 2.52% means that, on average, the deviation of E[RT] predicted by the eLD-model from the empirical means makes 2.52% of the empirical values. For both means and standard deviations, the models can be ordered according to goodness-of-fit, eCD>eLD>gLD. However the differences are so small that no model can be rejected. For the means, each model yields values of MSRD less than 5%, obviously a very good fit. If 5% is acceptable for means, then the MSRD values provided by the models for standard deviations are comparably good.*

The small differences in values of fit make one wonder whether the obtained ordering of the models --eCD>eLD>gLD-- is replicable. In other words, can one expect to get the same ordering if the experiment is repeated? The results of the auxiliary experiment, with three different dot densities, suggest that the answer should be negative. The number of velocity pairs used in this experiment was rather small, and only one value of V_1 was paired with V_0 equal to 1, 2, and 8 deg/s. However the remaining three values of V_0, 0, 4, and 16 deg/s, were paired with more than one value of V_1 each, and these pairs can be used for model fitting. The results are presented in the bottom of Table 1.

* This can be shown as follows. The experiment was carried out in three blocks each containing about 50 trials per $\langle V_0, V_1 \rangle$ pair. The MSRD of the three sets of within-block means from the set of grand means is 4.19% for RS and 3.37% for JF, both values below 5%. One can conclude that the three blocks of measurements per condition are mutually consistent, and that their consistency is comparable with the MSRDs for E[RT] versus mean. Then it is natural to compare the MSRDs for S[RT] versus st. dev. with the level of consistency of the within-block st. dev.s. The latter is calculated as MSRD of the three sets of within-block st. dev.s from the set of grand st. dev.s. The level of consistency is 30.11% for RS and 56.76% for JF, which is well above the MSRDs provided by the three models. This informal consideration makes it obvious that the variability of st. dev.s is of a greater order of magnitude than the variability of means. If 5% is acceptance level for means, then 25% for standard deviations seems to be a very conservative estimate.

encouraged us to use a statistic whose theoretical distribution was not known. This statistic is the relative deviation, |predicted-observed|/observed, which expresses differences between predicted and observed values as a percentage of the observed value. This dimensionless measure can be used for both means and standard deviations, and seems to be a natural choice for inherently positive data, such as RTs. For the central tendency of relative deviations we used Mean Squared Relative Deviation, MSRD:

$$MSRD = \{\Sigma[\,(\text{predicted} - \text{observed})/\text{observed}\,]^2/n\}^{1/2} * 100\% \qquad [10]$$

where summation is over all data points, that is for all n pairs, $\langle V_0, V_1 \rangle$. "Predicted" and "observed" should be replaced with either E[RT] and mean, or S[RT] and empirical standard deviation.

Theoretical predictions of the eLD-model are shown in Figures 3-6 by solid lines. The chosen format of the x-axis makes the predictions linear for mean RTs, and, in the range of velocity differences used, almost linear for standard deviations. The values of free parameters at which the minimum MSRD is achieved are given for all three models in the legends to Figures 3-6. In order not to impair readability we did not present the theoretical predictions of the two other models in the same plots, and presenting them separately would have taken too much space. The reason for singling out the eLD-model will be explained below. However it is not based on the values of minimum MSRD achieved by each model, as one can see from Table 1.

[Insert Table 1 about here]

When RTs were averaged over the three dot densities the ordering of the models was eLD>eCD>gLD. If the RTs corresponding to different numbers of dots were fitted separately, so that Δ and C are functions of both V_0 and dot density, then the resulting ordering was eLD>gLD>eCD. As we see, there is no consistent pattern in ordering of the models according to goodness-of-fit. In addition, the small differences between the MSRD values are at least in part due to the technical fact that we use precise computational formulas for the eCD-model, but only approximate formulas for the variants of the LD-model.

DIRECTION CHANGES VS. SPEED CHANGES. Figures 3-6 corroborate the ordinal characteristic of the data that we mentioned earlier: there were no qualitative differences between responses to 180^0 reversal of direction, on one hand, and responses to change in speed only, on the other. First, we verified that the fitted values of parameters were determined mainly by the unidirectional velocity pairs, rather than by the pairs with direction reversal; ignoring data involving a change in direction, and fitting models only to speed change data, produces very little change in the optimal values of models' parameters. This is not surprising since there were six times as many unidirectional velocity pairs as those with direction reversal. If RTs to direction reversals formed a qualitatively separate group they would deviate from predicted values more than do the RTs to speed change. This obviously is not the case.

The homogeneity of data, particularly the homogeneity of data for both speed changes and direction reversals, bears on the the general problem of velocity encoding in the visual system. However the data's homogeneity has an additional meaning within the framework of the LD-models. Unlike the case of unidirectional speed changes, direction reversals cause any dot to pass twice over each spatial position along its trajectory. For spatial positions near the turn point this retracing leads to some luminance blur that could limit the applicability of the formulas based on kinematic function only (see the description of the LD-model above). The homogeneity of the data shows that the amount of blur in direction reversals was negligibly small.

BEST-FITTING PARAMETER VALUES. Since none of the models could be dismissed on the grounds of poor fit, we gave extra attention to the plausibility of the optimal values of the models' parameters.

The estimates of the time-dimensioned parameters, $E[t_R]$ and $S[t_R]$, are given in the legends to Figures 3-6. For the eLD-model these values are shown as the intercept points of the vertical axes with the theoretical curves (corresponding to infinitely large velocity difference, or zero closeness). The estimates of t_R given by the eCD-model, 214.5 ± 25.5 ms (RWS) and 209 ± 26.5 ms (JF), seem somewhat too high for residual times.** They are considerably higher than values reported for simple RT to long large high-intensity light flashes (Teichner and Krebs, 1972).

Estimates for the displacement-dimensioned parameters, C and Δ, are also given in the legends to Figures 3-6. Note that different measures of central tendency and variability were used for different models (see Table A1 in Appendix 1). For both means and standard deviation the greater the value of the displacement-dimensioned parameters, the greater the predicted rate of data decrease as the velocity difference increases.

Of primary interest for us here are the values corresponding to $V_0=0$, the particular case when the change of velocity is the onset of a uniform motion. If

** Here, for compactness of presentation, we use the format $E[t_R]\pm S[t_R]$. This should not be confused with anything like "confidence intervals" for $E[t_R]$. The $E[t_R]$ and $S[t_R]$ are independent estimates of two different parameters of a hypothetical distribution.

and only if the Proposition of Identity holds, RTs to motion onset can be considered as a particular paradigm of motion detection. Therefore by analyzing the values of $C(0)$ and $\Delta(0)$ one can find out whether a particular model is consistent with the Proposition of Identity, *i.e.* whether these values are close to estimates of C and Δ derived from experiments on foveal absolute motion detectability. Recollect that $[C(0)]^2$ in the LD-model is the critical value of the local dispersion (formula 2) at which a target is judged as moving. The parameter $\Delta(0)$ in the CD-model (formula 4) is the critical distance that has to be traversed by a target to be judged as moving. In discussing detectability, the argument (0) in $C(0)$ and $\Delta(0)$ is redundant and can be dropped.

In order to compare the values of C and Δ directly, one can bring them to a "common denominator" by expressing them in values of amplitude thresholds for a fixed kinematic function. The simplest choice of the kinematic function is the instantaneous shift of position. As it was stated in the Introduction, if the CD-model can be related to detectability at all, then the amplitude threshold for instantaneous shift of position gives the most precise estimate of the critical displacement. In other words, the *equivalent threshold amplitude of instantaneous shift* for Δ (if the CD-model holds) is Δ itself. It can be shown that the *equivalent threshold amplitude of instantaneous shift* for C (if the LD-model holds) is equal to $C(6T/\tau)^{1/2} = 3.464C$ (since $T/\tau = 2$).

For their own data and from their reanalysis of others' data, Dzhafarov and Allik obtained values of C that fell between 0.1 - 0.7 min of arc. This can be considered a realistic confidence interval for $E[C]$. However in the analysis underlying these estimates -- for kinematic thresholds, psychometric functions, or reaction times -- C has been treated as a deterministic constant. The proposition that the estimated deterministic C-values are close to $E[C]$ is, strictly speaking, only a hypothesis. Therefore, in order to be absolutely sure, we will set more conservative interval 0.07 - 1.0 min of arc. It is hardly conceivable that E[C] for foveal absolute motion detection ought to fall outside these very generous boundaries. Indeed, the threshold amplitudes of instantaneous shift equivalent to these values are 0.25 - 3.5 min of arc, and the reported values of absolute shift thresholds lie well within these boundaries (Legge & Campbell, 1981). Obviously, these boundaries, 0.25 - 3.5 min of arc, should be considered also as a conservative interval for possible values of Δ.

Now, if the estimates of Δ and C are obtained from the reaction time rather than threshold experiments, then the Proposition of Identity can be judged to hold only if a central tendency of C and Δ falls between the established boundaries. This is what we are going to check for the values of C(0) and Δ(0) estimated from our present experiment.

The conservatism of our estimated boundaries for C and Δ makes the precise choice of the measure of central tendency for them rather unimportant: shift amplitudes of 0.24 min and 3.5 min certainly correspond to detection probabilities close to 0 and 1, respectively. However for direct comparison one should use a same measure of central tendency for both C and Δ. The measures estimated in our present analysis differ: it is $E[\Delta(0)]$ in the eCD-model, but it is $E[C(0)^{1/2}]^2$ in both versions of gLD-model. Fortunately we can easily avoid comparing moments of different types, since together with $E[C(0)^{1/2}]^2$ we get an independent estimation of $S[C(0)^{1/2}]^2$, and the sum of the two values should equal $E[C(0)]$.

[Insert Figure 8 about here]

In Figure 8 the value of $E[\Delta(0)]$ is plotted along with the estimations of $E[C(0)]$ derived from the eLD-model and gLD-model, multiplied by 3.464 to represent the equivalent shift thresholds. The figure illustrates the fact that $E[\Delta(0)]$ estimated in the eCD-model grossly exceeds the very conservative upper limit we have set: estimates are 6.62 min (RWS) and 5.11 min (JF). In contrast, derived from the eLD-model, 1.92 min (RWS) and 1.13 (JF) not only fall between the conservative margins, but are also well within the more "realistic" interval 0.35 - 2.4 min of arc. The obvious conclusion is that the considered variant of the LD-model generalization is nicely consistent with the Proposition of Identity, whereas the generalization of the CD-model is grossly inconsistent with it. In other words, if one accepts the eCD-model one must also accept the idea that the decision to react to the onset of motion is always made considerably after motion is actually detected.

The interpretation of the gLD-model is somewhat less certain. Although the two estimates, 2.76 min (RWS) and 1.75 min (JF), are within our conservative boundaries, the former value exceeds the "realistic" (with most probability also rather conservative) upper margin we have set. In combination with the fact that the fit provided by the gLD-model is slightly worse than that of the eLD-model, this makes the latter more preferable.

One may wonder why estimates of $C(0)$ given by the gLD-model and eLD-model differ when the two models are coincident at $V_0=0$, where the models converge onto the original form of LD-model. The reason is that the two models, gLD and eLD, are fitted to the entire set of data, and that the common parameter t_R makes the fit for different V_0-values interdependent.

DISCUSSION

COMPARISON OF THE MODELS. It was disappointing not to be able to choose among the three models on the basis of their fits to data. However there are other grounds for making a choice. For one thing, the CD-model is clearly not consistent with the Proposition of Identity, the assumption that observers use the same criterion in reaction time and detection experiments. Therefore accepting the eCD-model for RTs to velocity change (including motion onset/offset) would uncouple RT experiments from detection experiments. Such an uncoupling would pose some difficult questions: (1) why should different criteria control the observer's decision in the two types of experiments? (2) why would an observer in a reaction time experiment not respond as soon as the motion had been detected, particularly since the instructions clearly encourage such behavior?

None of these difficulties attends the LD-model. It provides a unified framework for both detectability and RT data, and justifies considering the latter as a special case of the former. Although, there is no logical necessity for the Proposition of Identity, in the absence of other factors Occam's razor compels a preference for a model in which a single principle gives rise to various forms of motion detection.

Comparison of the two versions of the LD-model favors the eLD-version over the gLD-version. For one thing, the eLD-model fits data slightly better (see Table 1). Second, it is in better agreement with the Proposition of Identity: the estimation of E[C] for RWS is slightly over the "realistic" upper boundary we had set. In addition, the eLD-model can be computationally simplified with a better precision. However, the superiority of the eLD-version should be taken with a reservation: the imprecision of the computational formulas for the gLD-model could itself have been responsible for the latter's worse performance.

NETWORKS OF BILOCAL CORRELATORS. In the rest of the paper we will consider the problem of realizability of the LD by a system of biologically plausible mechanisms. First, we will discuss this problem for the original motion detection model, then for the modifications of the eLD type. The LD-model for motion detection has been formulated as a highly specialized algorithm: it is applicable only if the moving stimulus, a spatio-temporal distribution of luminance, is represented by a single kinematic function defined at every moment. The problem of how the kinematic function is extracted from the stimulus flow-field is closely related to the general issue of the detection of non-rigid motion. Both questions are beyond the scope of this paper. However it is easy to see that a natural step toward solution of these problems is to realize the LD algorithm by the mass activation of more primitive and more universal mechanisms. The response of such a system to a rigidly moving pattern should be equal to the value of LD, but the system should perform computations over any spatio-temporal luminance distribution, however deviant from rigid motion.

One such system is suggested by the computational algorithm shown in Figure 2, and by the form in which moving variance is represented in equation [2]. Variance of a set of numbers is the mean squared deviation of the numbers from their mean, but it is also the mean squared pair-wise distance between the numbers themselves. Thus, in Figure 2, the variance of spatial positions within the travelling τ-window is proportional to the sum of all squared pair-wise distances between the spatial positions within the window. This suggests the idea that the variance could be provided by a pool of mechanisms each tuned to a particular temporal and spatial distance. The output of such a mechanism should be proportional to the squared spatial distance to which it is tuned.

It is not difficult to see in these mechanisms a variant of the widely accepted idea that *bilocal correlators* are the elementary units of visual motion encoding (Reichardt, 1961; Barlow and Levick 1963; van Doorn and Koenderink, 1982a,b; van de Grind, Koenderink, van Doorn, 1983). A bilocal correlator (Figure 9) consists of two units that sense the luminance profiles falling within two identical receptive regions separated by a distance Δs. The responses to the two luminance profiles are transmitted with a relative delay Δt into a comparator that performs a matching operation equivalent to a point-to-point correlation. For simplicity we will assume that a bilocal correlator is completely specified by Δt and the locations, s_1 and s_2, of its receiving regions, as if all bilocal correlators had the same size and the same sensitivity profile. This simplification will not affect the generality of our analysis, since it will be confined to rigid motion only. Note that Δs is the absolute value of the 2-D vector s_2-s_1 (or, if we consider only one-dimensional motion, s_2-s_1 is a signed number).

At a moment t, the output of a bilocal correlator, $\langle\Delta t, s_1, s_2\rangle$, is maximal if a same luminance profile occupied locations s_1 and s_2 at times $t-\Delta t$ and t. With a threshold device connected to the comparator (see Figure 9) the mechanism becomes a detector with a Boolean output (0 or 1): it "fires" at time t if and only if the patterns at $(t-\Delta t, s_1)$ and (t, s_2) match. In order to make the bilocal correlators compute a moving variance one has to make two additional assumptions. First, the output of a mechanism $\langle\Delta t, s_1, s_2\rangle$ should be multiplied by $\Delta s^2 = |s_2-s_1|^2$ (Figure 10, upper panel). Second, this output should last for τ-

Δt (Figure 10, lower panel).

The first assumption, multiplication by Δs^2, can be thought of in many "technical" variants. Thus, it could mean a straightforward amplification of the Boolean output, or it could mean that the number of identical mechanisms $\langle \Delta t, s_1, s_2 \rangle$ with Boolean outputs is an integer approximation of Δs^2. It might even have no structural meaning at all: since the output of any bilocal correlator is on a "labeled line", it can be "taken with appropriate weight" on a subsequent processing stage. Whatever the technical aspect of the multiplication, its *functional* meaning is the following. In a network of bilocal mechanisms designed for detection of motion, the detection of larger displacements conveys more evidence for motion than the detection of smaller ones. Therefore responses of the bilocal correlators should be taken with weights monotonically related to their spatial span, Δs. Squaring is a particular choice of such a monotonic function.

The second assumption, above, means that the total duration of the mechanism's cycle of activity, starting with activation of its first sensing unit, is τ: the cycle is comprised of the transmission time, Δt, and the output time, $\tau - \Delta t$. It follows that the maximum value of Δt a bilocal mechanism can have is τ, with instantaneous output. Since a new cycle of activity of any mechanism is initiated at every moment of time, the assumption should be complemented by some rules of interaction of subsequent cycles. For simplicity we assume no-interaction: the images of subsequent luminance profiles are transmitted to the comparator independently, and the overlapping outputs summed.

The summary output of a pool of the described mechanisms at any moment t will be proportional to moving variance of the kinematic function, provided all triads $\langle \Delta t, s_1, s_2 \rangle$, $\Delta t < \tau$, are represented in the pool. Of course, in a real network the representation can be only provided by a finite set of mechanisms with overlapping spatial and temporal tuning. Therefore the proportionality of the network's output to the moving variance of the kinematic function can only be approximate.

Moving variance is only first step in the computational algorithm shown in Figure 2. To obtain LD one has to "smooth" the moving variance function by the T-length moving average operator. The realization of this final stage in terms of bilocal mechanisms is straightforward. Outputs of all the mechanisms should be assumed to feed into a leaky integrator, or "stack" of temporal span T (Figure 10, upper panel). Recall that the operation of averaging provides an estimation of the magnitude of the moving variance function. Thus if T is zero then the magnitude of the function will be the maximal single value of the moving variance; if T is infinitely large then the magnitude is the grand mean of all variance values. The actual value of T lies between these two poles. The output of the T-length "stack" at every moment t is proportional to the LD-value given by formula [2]. Namely, it is equal to $LD(t)T\tau^2$, and in decision rules postulated for threshold setting and reaction initiation it should exceed the critical level $C^2T\tau^2$.

In our description of bilocal correlators we have not specified whether the receiving areas of a correlator are defined in retinal or stimulus-plane coordinates. Either can be true. One could even assume that motion is processed on two levels: a lower-level retina-bound network of bilocal correlators, and a higher-level network with a built-in compensation for eye movements. The question is which of these networks is associated with motion detection. In most motion detection paradigms eye movements are negligible, so neither possibility can be rejected. Therefore, in the context of this paper, we will consider implications for velocity change detection associated with each of these possibilities: what additional assumptions should be made, or how the network of bilocal mechanisms can be modified, to realize the eLD-model for detection of velocity changes.

If motion detection is defined in retinal coordinates, then the simplest hypothesis seems to be following.*** Since no fixation point was provided in our experiments, and the duration of the first phase of motion was relatively long (between 1 and 2 s), the observers certainly reached the smooth-pursuit stage of eye movement during this phase. Therefore, as velocity changes from V_0 to V_1, the retinal velocity changes from 0 to $|V_1-V_0|$, precisely the equivalence postulated in the eLD-model. One has to make additional assumptions to explain the increase of the critical level C as V_0 increases from 4 to 16 deg/s. One could assume that tracking of faster motions is associated with a higher level of "noise", or "residual activity" in the network of bilocal mechanisms, which (applying a standard signal-to-noise analysis) should be compensated for by adoption of a higher criterion level. The higher level of residual activity when tracking faster motions could be attributed to any or all of the following factors: first, the initial activity in the network, before a catching-up-with-V_0 saccade, is higher for faster motions; second, tracking could be less smooth for faster motions; finally, the average time of uninterrupted tracking decreases as motion velocity increases. Indeed, if tracking starts in the center of our 16 deg aperture, then for 8 and 16 deg/s velocities the eye would have to return to the center and start over again 1-2 times and 2-4 times, respectively. No returns would be necessary for velocities of 0-4 deg/s, so any residual activity following the initial catching-up-with-V_0 saccade would have more time to diminish.

*** The authors are indebted to Joseph Malpeli for substantial contribution into this hypothesis.

If motion detection is defined in stimulus-plane rather than retinal coordinates, then some form of the "adaptation" process should replace the physical zeroing of forespeed in the previous hypothesis. The required process can be provided by a re-calibration of the weights, or amplification coefficients, attached to the Boolean outputs of the bilocal correlators. Namely, at the second phase of motion, V_1, the output of any bilocal mechanism $\langle\Delta t, s_1, s_2\rangle$, instead of being multiplied by $\Delta s^2 = |s_2 - s_1|^2$, should be multiplied by $|(s_2 - s_1) - V_0\Delta t|^2$. Let us consider in more detail the process by which adjustment of weights might be achieved. During the first phase of the two-phase motion $\langle V_0, V_1\rangle$ in every subset of the bilocal mechanisms corresponding to a given Δt the activated mechanisms in the network will group around the elements $\langle\Delta t, s, s+V_0\Delta t\rangle$ (provided that the subset is activated at all, *i.e* if the motion has lasted for more than Δt). This excitation pattern becomes stabilized after a time close to τ, and the task is to detect the change in this pattern. This goal is achieved by the re-calibration of the system of weights attached to the mechanisms, so that after the period τ the network would not respond until the excitation pattern changes. The re-calibration is mathematically equivalent to subtracting the spatial span $V_0\Delta t$ of the excited mechanisms from spatial spans of all mechanisms with a given temporal span Δt. After that, as long as the first phase of motion lasts, the reorganized system will be silent: the responses of the excited mechanisms will be multiplied by $|(s+V_0\Delta t) - s - V_0\Delta t|^2 = 0$. As soon as the velocity changes to V_1, the now-reorganized system will respond like the original system would have responded to $V_1 - V_0$: the outputs of the excited mechanisms $\langle\Delta t, s, s+V_1\Delta t\rangle$ will be multiplied by $|(s+V_1\Delta t) - s - V_0\Delta t|^2 = |(V_1 - V_0)\Delta t|^2$. The hypothetical process of re-calibration, providing a transient character of motion detection network activity, could be referred to as "self-inhibition".

To understand why the reorganization of weights also affects the critical level C, one could again assume that silencing of the network is only relative, and that a "residual activity" is higher for faster motions. One could even repeat one of the arguments suggested in the retina-bound-network hypothesis: that higher residual activity is due to the higher initial activity realizing detection of the first phase of motion. Also, the necessity to restart tracking after encountering aperture border could be associated with a re-activation of the network even if defined in stimulus coordinates. Alternatively, or in addition, one could assume that spatial tuning characteristics of bilocal mechanisms overlap, and that the degree of overlap increases with Δs. Consider the set of bilocal correlators with a given span Δt. Suppose that during the V_0-phase three groups of mechanisms were activated, with peak spatial tuning to $V_0\Delta t$, $V_0\Delta t+\varepsilon$, and $V_0\Delta t-\varepsilon$. The assumption we have made above means that ε is greater for greater $V_0\Delta t$, and thereby for greater V_0. One of the values, $V_0\Delta t$, $V_0\Delta t+\varepsilon$, or $V_0\Delta t-\varepsilon$ should be chosen to serve as an effective zero in the modified system of weights attached to the mechanisms with the temporal span Δt. At the present level of analysis it is immaterial whether the effective zero is chosen at random amidst the activated units, or whether there is a mechanism determining the "central" value $V_0\Delta t$ more precisely. Whatever the rule, it is clear that the "silencing" of the network at the end of the V_0-phase, after the weights have been re-calibrated, is only relative. For example, if $V_0\Delta t$ operated as an effective zero point, then the responses of the mechanisms tuned to spatial shifts $V_0\Delta t+\varepsilon$ and $V_0\Delta t-\varepsilon$ will each be taken with the weight $|(V_0\Delta t+\varepsilon)-V_0\Delta t|^2 = \varepsilon^2$. Applying a standard signal-to-noise analysis, greater values of ε will require the adoption of higher critical levels.

CONCLUSION. We conclude this paper with a brief recapitulation of the main results. First, a modified variant of the LD-model accounts for the RTs to velocity changes $\langle V_0, V_1 \rangle$. The essence of this modified variant is the application of the original LD-model to the detection of motion onset in $\langle 0, V_1 - V_0 \rangle$, with the critical level C being a (non-strictly) increasing function of V_0. Second, at $V_0=0$, where the modified and the original versions of the model logically coincide, the estimated value of C was found to be in a good agreement with the estimates obtained from other motion detectability experiments. Third, the changes in speed and direction are treated in the same way. In both cases, the perceptual response seems to depend upon the algebraic difference between $V_1 - V_0$. Finally, both the original and the modified versions of the LD-model can be realized by mass activation of a network of bilocal mechanisms.

Some of the characteristics we have attributed to these bilocal mechanisms do not seem to have obvious analogues in known physiological structures. The long duration of the mechanisms' activity, about 0.5 s, suggests that the analogues should be sought in the neuronal *circuitry* rather than in single neurons. However physiological considerations do not seem to be most imminent problem at present. Many questions remain to be answered in a purely psychophysical plane. Thus, it is not clear how the described network can provide the concordant shift of τ and C as the detection changes from absolute to relative motion (Dzhafarov and Allik, 1984). Also, it remains to be found out, whether the network can account for the detection of non-rigid planar motion. This seems to be a very important line for future analysis, which should show whether the model can indeed be considered as a good generalization of the original algorithm for local dispersion.

Speaking specifically about the problem of RTs to velocity changes, an important remaining problem is to experimentally test the hypothesis of eye movements against the hypothesis of re-calibration of weights. Another obvious continuation of the present work would be to use two dimensional velocity pairs, i.e. pair of V_0 and V_1 that differ only in the orientation of their motions. The eLD-model, described in this paper, can be applied without modification to this situation if $|V_1-V_0|$ is understood to be the length of a vectorial difference, rather than as the absolute value of a scalar.

Acknowledgments. This work was partially supported by a grant from the U.S. Air Force Office of Scientific Research, AFOSR 85-0370. The authors are grateful to Joseph Malpeli for the critical reading of the manuscript.

APPENDIX

COMPUTATIONAL FORMULAS FOR eCD-MODEL, eLD-MODEL, AND gLD-model

Formulas for $E[t_D]$ and $S[t_D]$ for the eCD-model can be derived from formula [5]:

$$\cdot[t_D(V_0,V_1)] = \cdot[\Delta(V_0)]/|V_1-V_0| \qquad \text{(A1)}$$

where the period denotes either of two moments, E and S. This formula together with the general equations [9] form the computational basis for the eCD-model-predictions.

The situation is more complicated with the two models that are based on LD-model. In formulas [6] and [8] there is no function of $C(V_0)$ on which t_D depends linearly. Strictly speaking, to deal with the problem we have to specify the exact form of the 8 distributions of $C(V_0)$s, V_0 = 0, 1, 2, 4, 8, and 16 deg/s. However such an analysis would add more free parameters and make the LD-model-based versions incomparable with the simple application of the eCD-model.

Fortunately there is a way to avoid such an awkward analysis. We can assume that the decision time, t_D, is considerably smaller than τ (0.5s). Then in formulas [6] and [8] all the summands except those containing the lowest power of the fraction t_D/τ can be omitted. This assumption gives us approximate formulas in which t_D depends linearly on some (nonlinear) function of $C(V_0)$. Now the formulas for the moments can be easily derived. For the eLD-model we have:

$$.[t_D(V_0,V_1)] = .[C(V_0)^{1/2}](12T\tau)^{1/4}/(|V_1-V_0|)^{1/2} \qquad \text{(A2)}$$

For the gLD-model we have:

$$.[tD(V_0,V_1)] = .[C(0)^{1/2}](12T\tau)^{1/4}/|V1|^{1/2} \qquad \text{if } V_0 = 0$$

$$\text{(A3)}$$

$$.[tD(V_0,V_1)] = .[C(VO)^{2/3}](6T)^{1/3}/(|V_1-V_0|)^{1/3} \qquad \text{if otherwise}$$

Here again the period stands either for E or S, and the predictions for E[RT] and S[RT] are derived by combining the formulas with the general equations [9]. The values of T and τ in application of the formulas were put equal to 1s and 0.5s, respectively. The value of T/τ has been shown to equal 2 for all detection experiments, whereas the value of τ varied in the region 0.4 - 0.7s. The value 0.5s for present analysis was chosen simply as a "round" number. We have checked that change of τ value in the region 0.4 - 0.7s leads to only minor changes in predicted values. All three models have the same two time-dimensioned parameters, $E[t_R]$ and $S[t_R]$. The following table summarizes the sets of the models' distance-dimensioned parameters.

[Insert Table A1 about here]

TABLE 1. MINIMUM MSRD VALUES

	mean RT			st. dev. of RT		
SUBJECT	eCD	eLD	gLD	eCD	eLD	gLD
RWS	3.57%	3.69%	4.80%	13.66%	13.67%	14.66%
JF	2.52%	2.67%	3.12%	18.92%	20.17%	20.88%
JLM*	2.63%	1.91%	2.66%			
JLM**	2.54%	2.12%	2.49%			

* auxiliary experiment, averaged over 3 dot densities

**auxiliary experiment, 3 dot densities fitted separately

TABLE A1. DISTANCE-DIMENSIONED PARAMETERS

MODEL	PARAMETER	CENTRAL TENDENCY	VARIABILITY
eCD-model	$\Delta(V_0)$	$E[\Delta(V_0)]$	$S[\Delta(V_0)]$
eLD-model	$C(V_0)$	$E[C(V_0)^{1/2}]^2$	$S[C(V_0)^{1/2}]^2$
gLD-model	$C(V_0)$	if $V_0=0$ $E[C(0)^{1/2}]^2$	$S[C(0)^{1/2}]^2$
		if $V_0=0$ $E[C(V_0)^{2/3}]^{3/2}$	$S[C(V0)^{2/3}]^{3/2}$

REFERENCES

Allik J. and Dzhafarov E. N. (1984) Reaction time to motion onset: Local dispersion model analysis. Vision Research, **24**, 99-101.

Ball K. and Sekuler R. (1980) Models of stimulus uncertainty in motion perception. Psychological Review, **87**, 435-469.

Barlow H. B. and Levick W. R. (1963) The mechanisms of directionally sensitive units in rabbit's retina. Journal of Physiology (London), **178**, 377-504.

Bonnet C. (1977) Visual motion detection models: Features and frequency filters. Perception, **6**, 491-500.

Bonnet C. (1982) Thresholds of motion perception. In: A. H. Wertheim, W. A. Wagenaar and H. W. Leibowitz (Eds) Tutorials on Motion Perception. New York: Plenum, 41-79.

Cohen R. L. and Bonnet C. (1972) Movement detection thresholds and stimulus duration. Perception & Psychophysics, **12**, 269-272.

Dzhafarov E. N. (1982) General model for visual motion detection. Studia Psychologica, **24**, 193-198.

Dzhafarov E. N. and Allik J. (1984) A general theory of motion detection. In: M. Rauk (Ed) Computational Models in Hearing and Vision. Tallin: Estonian Academy of Sciences, 77-84.

Dzhafarov E. N., Allik J. and Linde N. D. (1983) Detection of oscillatory movement. Voprosy Psikhologii, **3**, 90-96 (in Russian).

Dzhafarov E. N., Allik J., Linde N. D. and Plastolov V. K. (1981) Comparative analysis of frequency-amplitude threshold functions for real and apparent motion. Psikhologicheskii Zhurnal, **2**, 73-78 (in Russian).

Johnson C. A. and Leibowitz H. W. (1976) Velocity-time reciprocity in the perception of motion: Foveal and peripheral determinations. Vision Research, **16**, 177-180.

Legge G. E. and Campbell F. W. (1981) Displacement detection in human vision. Vision Research, **21**, 205-213.

Nakayama K. (1985) Biological image motion processing: A review. Vision Research, **25**, 625-660.

Reichardt W. (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: Rosenblith W. A. (Ed) Sensory Communications, Massachusetts: MIT Press, 303-317.

Teichner W. H. and Krebs M. J. (1972) Laws of the simple visual reaction time.

Psychological Review, **79**, 344-358

Tynan P.D. and Sekuler R. (1982) Motion processing in peripheral vision: Reaction time and perceived velocity. Vision Research, **22**, 61-68.

Van Doorn A. J., and Koenderink J. J. (1982a) Temporal properties of the visual detectability of moving spatial white noise. Experimental Brain Research, **45**, 179-188.

Van Doorn A. J., and Koenderink J. J. (1982b) Spatial properties of the visual detectability of moving spatial white noise. Experimental Brain Research, **45**, 189-195.

Van de Grind W. A., Koenderink J. J. and Doorn A. J. van (1983) Detection of coherent movement in peripherally viewed random-dot patterns. Journal of the Optical Society of America, **73**, 1674-1683.

FIGURE CAPTIONS

FIGURE 1. Display and types of kinematic functions used. Multiple-dot patterns like that shown in the upper left panel moved horizontally inside a 16 deg diameter circular aperture. The motion consisted of two phases, with constant velocities represented by the slopes of the straight lines in the panels a, a1, b, b1, and c. The two motions were either in the same direction (panels a, b), or in opposite directions (panel c). In the latter case the two phases had equal speeds. For unidirectional phases, the change in speed could be incremental (panel a) or decremental (panel b), including the cases of motion onset (panel a1) and offset (panel b1). See Procedure for details.

FIGURE 2. Schematic presentation of an algorithm equivalent to formula [2] of the LD-model. Right panel shows a complex kinematic function with temporal window of length τ travelling in time and computing the variance of spatial positions within it. Two positions of the τ-window are shown in the figure: $[t'-\tau, t']$ and $[t''-\tau, t'']$. The results of the computations form the moving variance function shown in the middle panel. Thus, the value of this function at moment t' is equal to the variance of spatial positions passed between the moments $t'-\tau$ and t'. The moving variance function is smoothed by travelling window of length T. This smoothing produces the LD-function

shown in the left panel. Two positions of the T-windows are shown in the figure: [t*-T,t*] and [t**-T,t**]. Thus the LD-value at the moment t* is equal to the mean value of the moving variance between t*-T and t*.

FIGURE 3. Mean RT versus "square-root-closeness" of V_1 to V_0, $|V_1 - V_0|^{-1/2}$. Subject JF. Every panel contains the mean RTs for all pairs $\langle V_0, V_1 \rangle$, but the means corresponding to one value of V_0 (given in insets) are "highlighted" (represented by squares), whereas the remaining values serve as a background (dots). Filled squares correspond to velocity increase ($V_1 > V_0$), empty squares with central dots correspond to velocity decrease ($V_1 < V_0$), crossed squares represent the direction reversal condition ($V_1 = -V_0$).
Solid lines are theoretical predictions of the eLD-model: $E[t_R]$ is equal to 180.5 ms (intercept with the vertical axis), central tendency of C (from panel 0 through 16) is equal to 0.28 - 0.31 - 0.37 -0.39 - 0.65 - 1.37 (min arc). These values correspond to the slopes of the solid lines.
Optimal parameters for the eCD-model: E[RT] = 209.0 ms; central tendency of Δ (from panel 0 through 16) is 5.11 - 5.13 - 6.30 - 7.99 - 13.66 - 28.21 (min arc).
Optimal parameters for the gLD-model: E[RT] = 163.0 ms; central tendency of C (from panel 0 through 16) is 0.45 - 1.08 -1.63 - 2.26 - 4.07 - 8.43 (min arc).
See Table A1 for the exact meaning of "central tendency".

FIGURE 4. Standard deviation of RT versus "square-root-closeness" of V_1 to V_0, $|V_1 - V_0|^{-1/2}$. Subject JF. Every panel contains the st. dev.s for all pairs $\langle V_0, V_1 \rangle$, but the st. dev.s corresponding to one value of V_0 (given in

insets) are "highlighted" (represented by squares), whereas the remaining values serve as a background (dots). Filled squares correspond to velocity increase ($V_1 > V_0$), empty squares with central dots correspond to velocity decrease ($V_1 < V_0$), crossed squares represent the direction reversal condition ($V_1 = -V_0$).

Solid lines are theoretical predictions of the eLD-model: $S[t_R]$ is equal to 22.0 ms (intercept with the vertical axis); variability of C (from panel 0 through 16) is equal to 0.046 - 0.067 - 0.084 - 0.058 - 0.118 - 0.237 (min arc). These values roughly correspond to the slopes of solid lines.

Optimal parameters for the eCD-model: S[RT] = 26.5 ms; variability of Δ (from panel 0 through 16) is 2.901 - 2.928 - 4.697 - 4.383 - 9.530 - 17.731 (min arc).

Optimal parameters for the gLD-model: S[RT] = 18.0 ms; variability of C (from panel 0 through 16) is 0.058 - 0.281 - 0.411 - 0.449 - 0.895 - 1.833 (min arc).

See Table A1 for the exact meaning of "variability".

FIGURE 5. Same as Figure 3, but for subject RWS.

Solid lines are theoretical predictions of the eLD-model: $E[t_R]$ is equal to 180.5 ms; central tendency of C (from panel 0 through 16) is equal to 0.49 - 0.49 - 0.39 - 0.50 - 0.88 - 1.64 (min arc).

Optimal parameters for the eCD-model: E[RT] = 214.5 ms; central tendency of Δ (from panel 0 through 16) is 6.62 - 7.12 - 6.19 - 8.24 - 15.43 - 28.13 (min arc).

Optimal parameters for the gLD-model: E[RT] = 162.0 ms; central tendency of C (from panel 0 through 16) is 0.72 - 1.4 - 1.7 - 2.66 -4.96 - 9.62 (min arc).

FIGURE 6. Same as Figure 4, but for subject RWS.

Solid lines are theoretical predictions of the eLD-model: $S[t_R]$ is equal to 19.5 ms, variability of C (from panel 0 through 16) is equal to 0.065 - 0.058 - 0.053 - 0.096 - 0.169 - 0.189 (min arc).

Optimal parameters for the eCD-model: S[RT] = 26.5 ms; variability of Δ (from panel 0 through 16) is 3.024 - 2.930 - 3.278 - 5.378 - 11.219 - 14.660 (min arc).

Optimal parameters for the gLD-model: S[RT] = 18.0 ms; variability of C (from panel 0 through 16) is 0.077 - 0.245 - 0.309 - 0.616 - 1.106 - 1.542 (min arc).

FIGURE 7. Results of the auxiliary experiment. Mean RTs for patterns with 50 and 100 dots at each value of $\langle V_0, V_1 \rangle$ are plotted against mean RTs with patterns of 200 dots for the same $\langle V_0, V_1 \rangle$.

FIGURE 8. Equivalent amplitude of instantaneous displacement corresponding to theoretical estimations of distance-dimensioned parameters, C and Δ, at $V_0 = 0$. If the Proposition of Identity holds, the equivalent amplitude should be equal to the minimal detectable amplitude for instantaneous displacement. Clear area in the figure corresponds to the range of realistic values for the amplitude threshold. Sparsely stippled area corresponds to the values that are beyond the realistic limits but still within the conservative boundaries set in this paper. Densely stippled area corresponds to the range that certainly cannot include possible values of the threshold amplitude. See Analysis for details.

FIGURE 9. Basic structure of a bilocal correlator. Two identical receptor areas centered at s_1 and s_2 feed into a matching device, or comparator. Information transmission from the s_1-area to the comparator takes by Δt longer than transmission from the s_2-area. Therefore the images of luminance profiles falling on the two areas at two moments separated by Δt reach the comparator simultaneously. The images are supposed to be analogues of spatial maps of excitation, and the comparator performs an operation analogous to a point-to-point correlation. If the value of this correlation exceeds a critical level set by the subsequent threshold device, the mechanism generates a signal. See Discussion for details.

FIGURE 10. Basic structure of a bilocal correlator that implements the LD-algorithm of Figure 2. Upper panel: the output of correlators is amplified proportionally to the squared value of their spatial spans, and is fed into a leaky integrator. The integrator acts as a stack whose memory span is T: at every moment t it adds the summary input to its content and "forgets" the input received at the moment t-T. Lower panel: the bilocal correlator's signal that is initiated by a given pair of luminance profiles, separated in time by Δt, lasts for $\tau-\Delta t$. So the total cycle of activity of any correlator takes a constant time, τ. See Discussion for details.

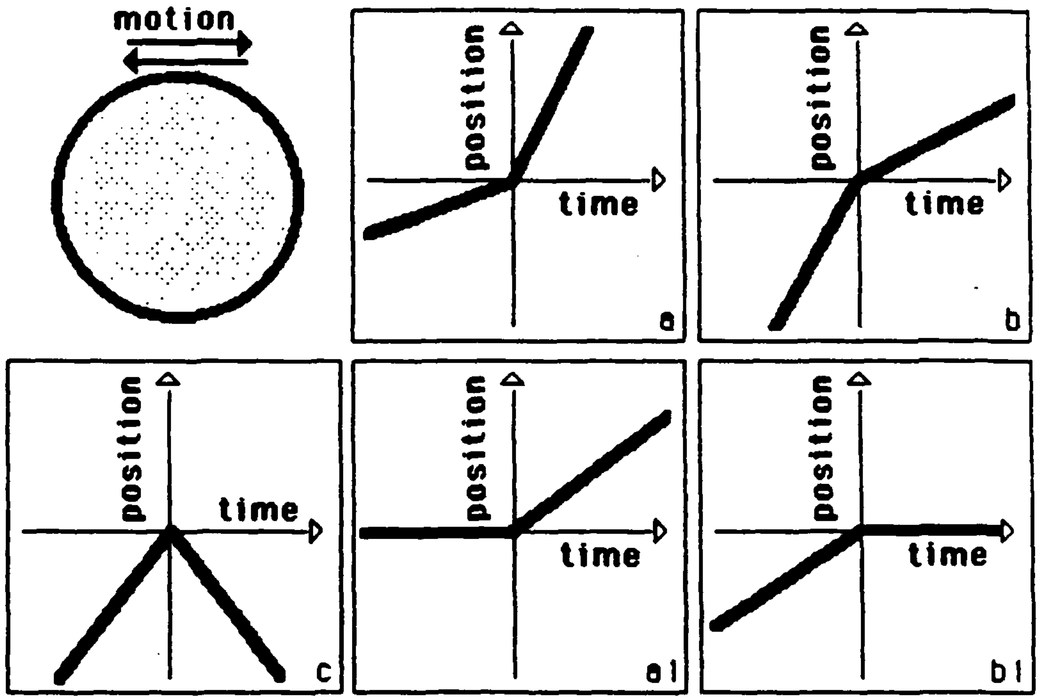
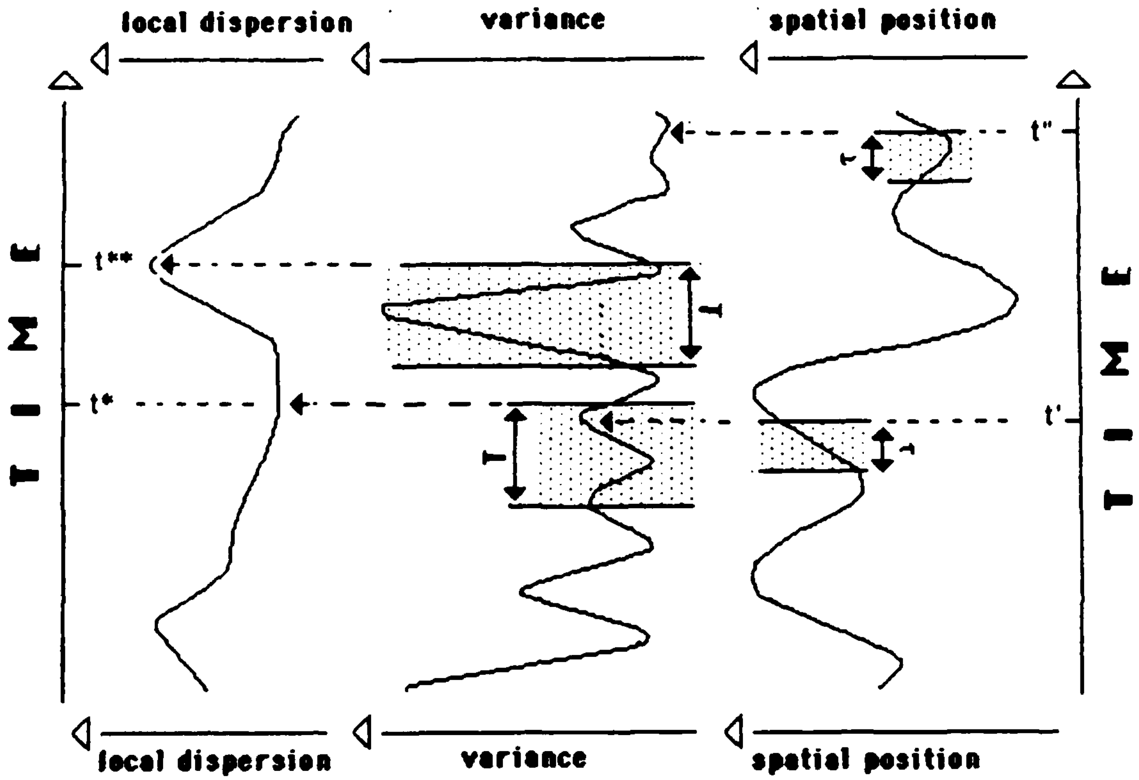
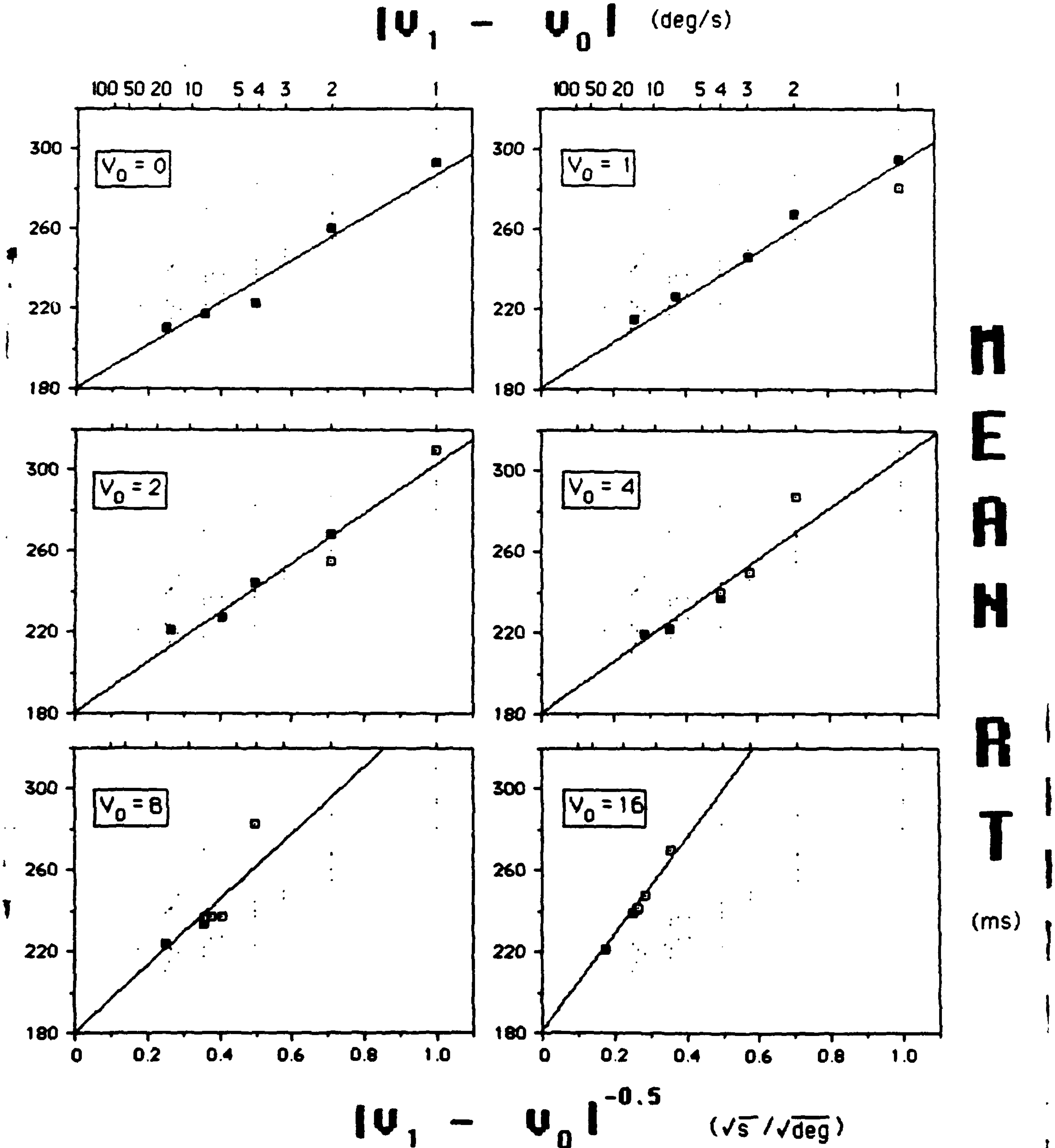


FIGURE 1



local dispersion
variance
spatial position
TIME
t**
t*
t"
t'
τ
T
TIME
local dispersion
variance
spatial position

FIGURE 2



|V1 - V0| (deg/s)
100 50 20 10 5 4 3 2 1
$V_0 = 0$
$V_0 = 1$
$V_0 = 2$
$V_0 = 4$
$V_0 = 8$
$V_0 = 16$
180
220
260
300
0 0.2 0.4 0.6 0.8 1.0
MEAN RT (ms)
$|V_1 - V_0|^{-0.5}$ ($\sqrt{s}/\sqrt{deg}$)

FIGURE 3

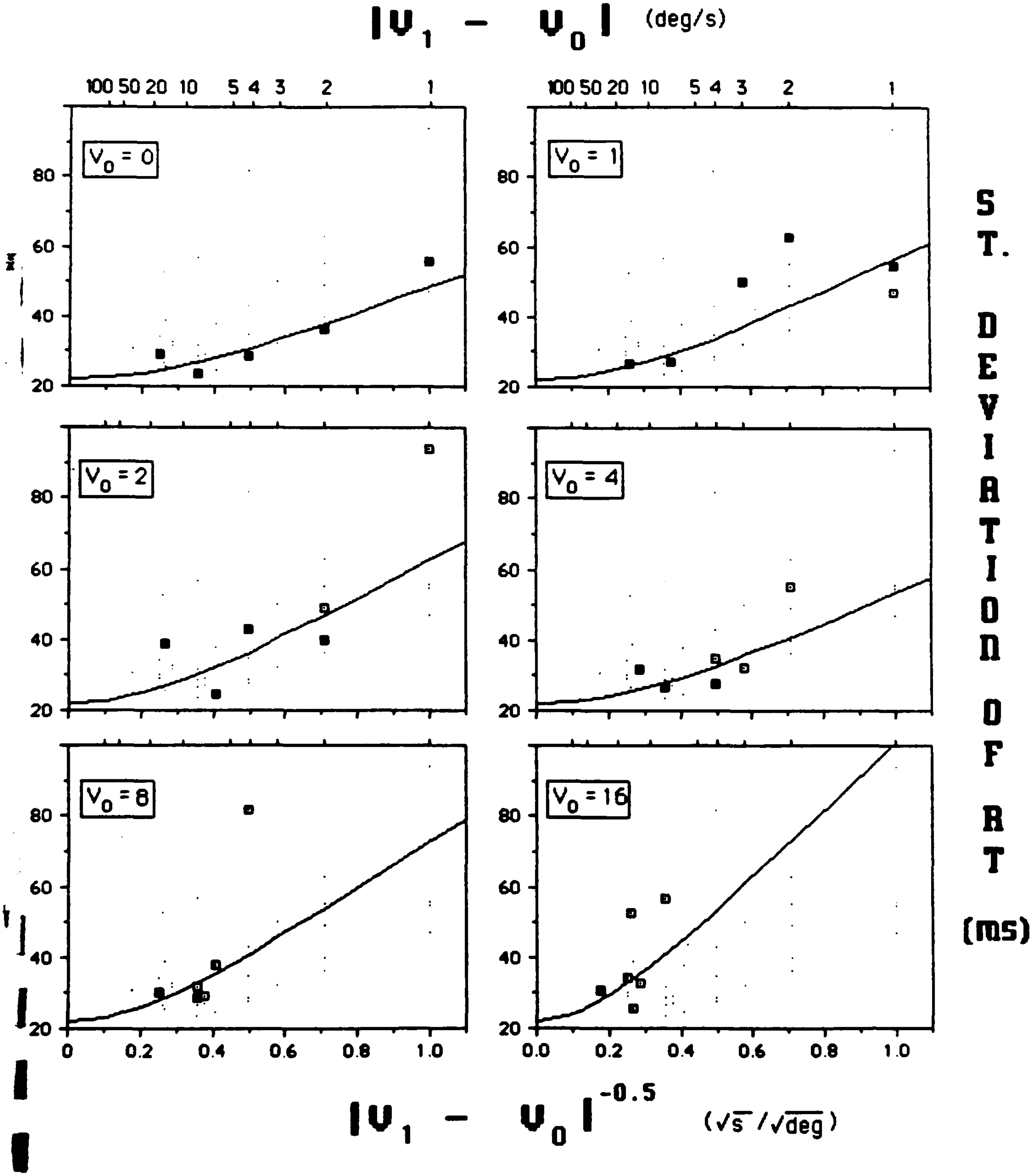

|V1 − V0| (deg/s)
100 50 20 10 5 4 3 2 1
V0 = 0
V0 = 1
V0 = 2
V0 = 4
V0 = 8
V0 = 16
20
40
60
80
0 0.2 0.4 0.6 0.8 1.0
0.0 0.2 0.4 0.6 0.8 1.0
|V1 − V0|^-0.5 (√s / √deg)
ST. DEVIATION OF RT (ms)

FIGURE 4

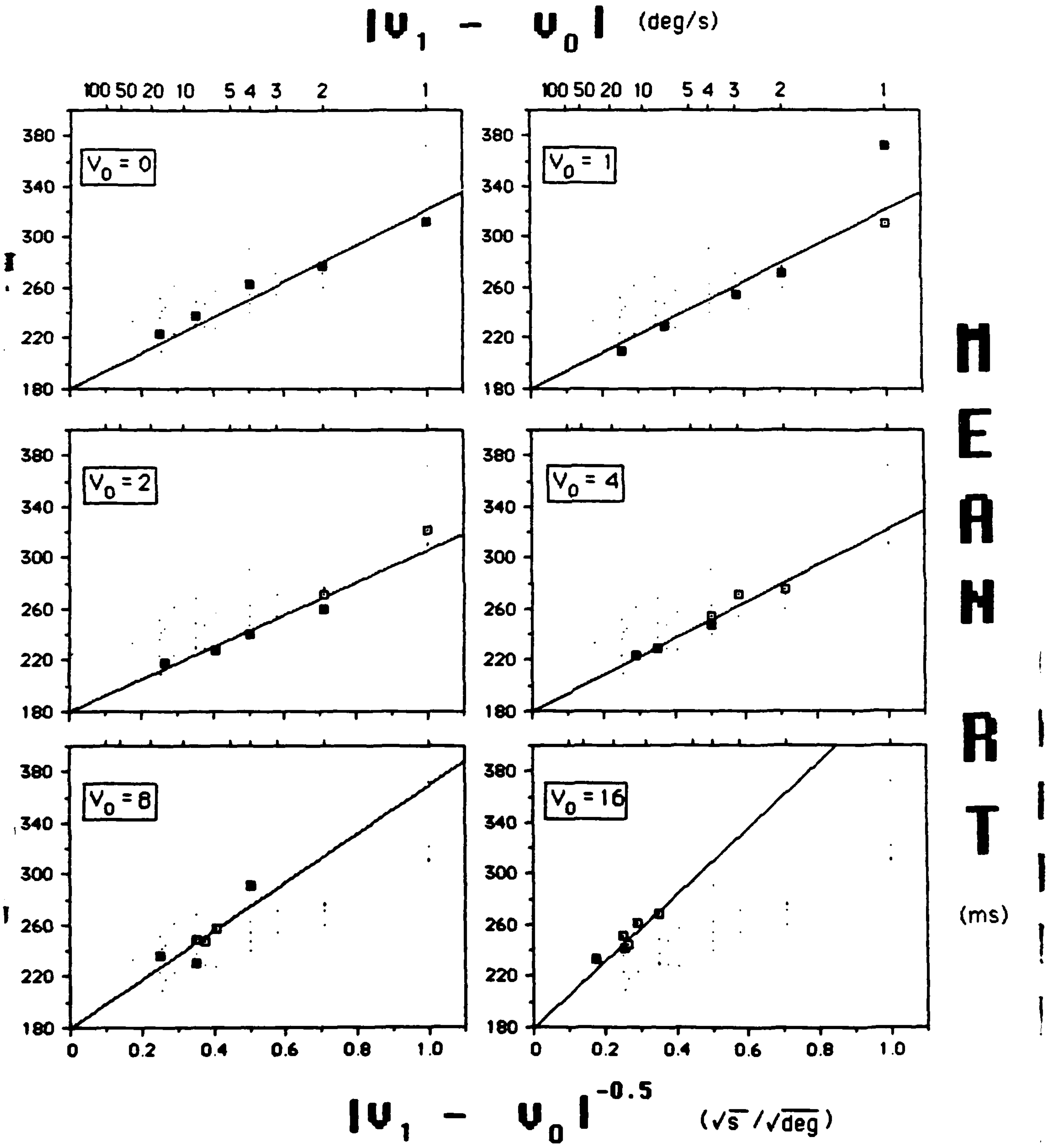
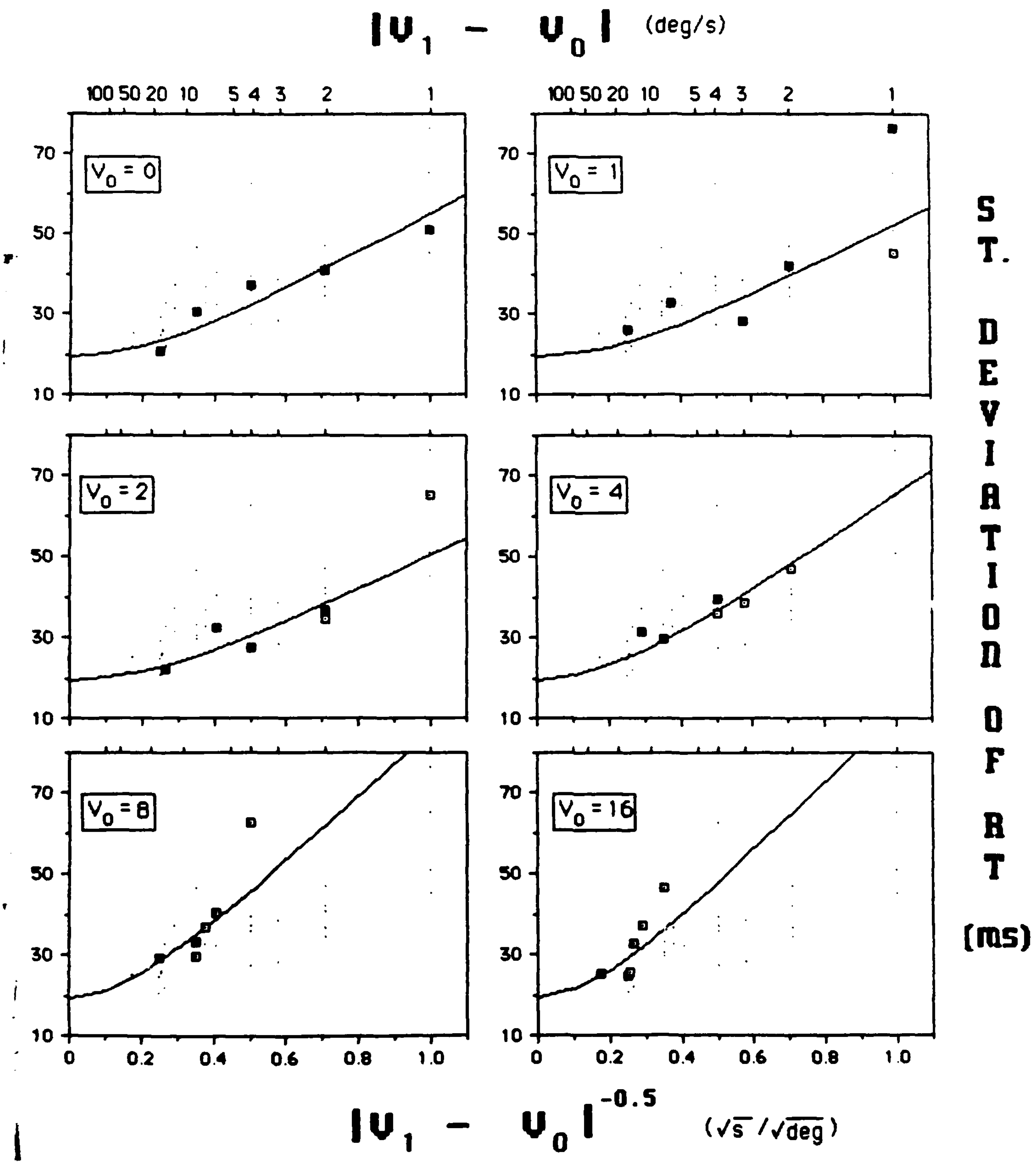

|V1 - V0| (deg/s)
100 50 20 10 5 4 3 2 1
V0 = 0
V0 = 1
V0 = 2
V0 = 4
V0 = 8
V0 = 16
380
340
300
260
220
180
0 0.2 0.4 0.6 0.8 1.0
|V1 - V0|^-0.5 (√s / √deg)
MEAN RT
(ms)

FIGURE 5



|V1 - V0| (deg/s)
100 50 20 10 5 4 3 2 1
V0 = 0
V0 = 1
V0 = 2
V0 = 4
V0 = 8
V0 = 16
10
30
50
70
0
0.2
0.4
0.6
0.8
1.0
ST. DEVIATION OF RT (ms)
|V1 - V0|^-0.5 (√s / √deg)

FIGURE 6

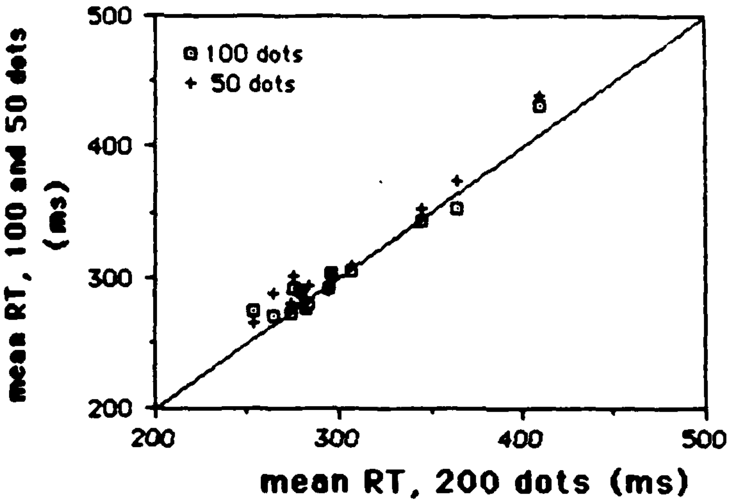
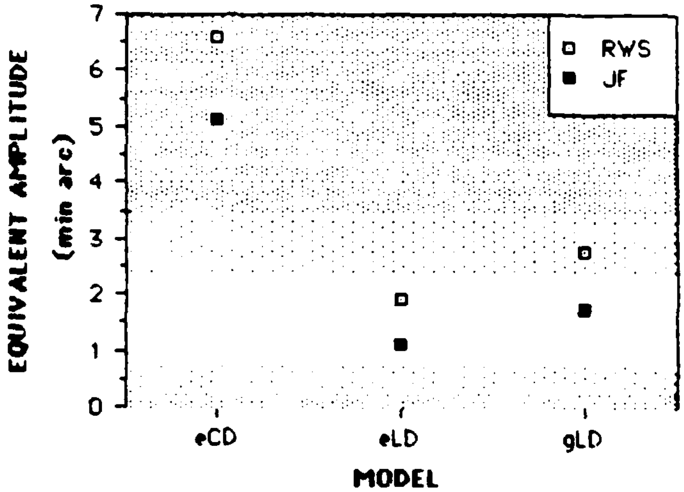

500
400
300
200
200
300
400
500
100 dots
50 dots
mean RT, 100 and 50 dots (ms)
mean RT, 200 dots (ms)

FIGURE 7



RWS
JF
EQUIVALENT AMPLITUDE
(min arc)
0
1
2
3
4
5
6
7
eCD
eLD
gLD
MODEL

FIGURE 8

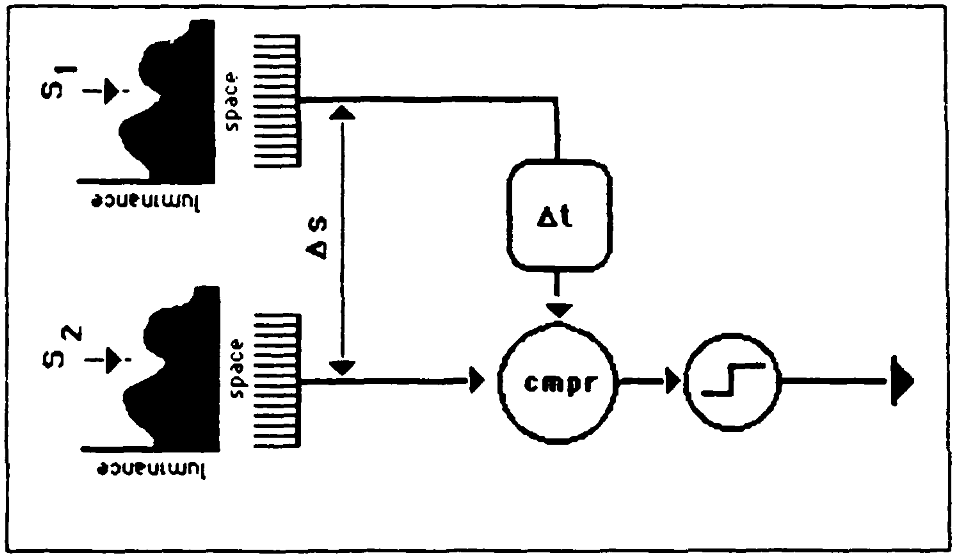


FIGURE 9

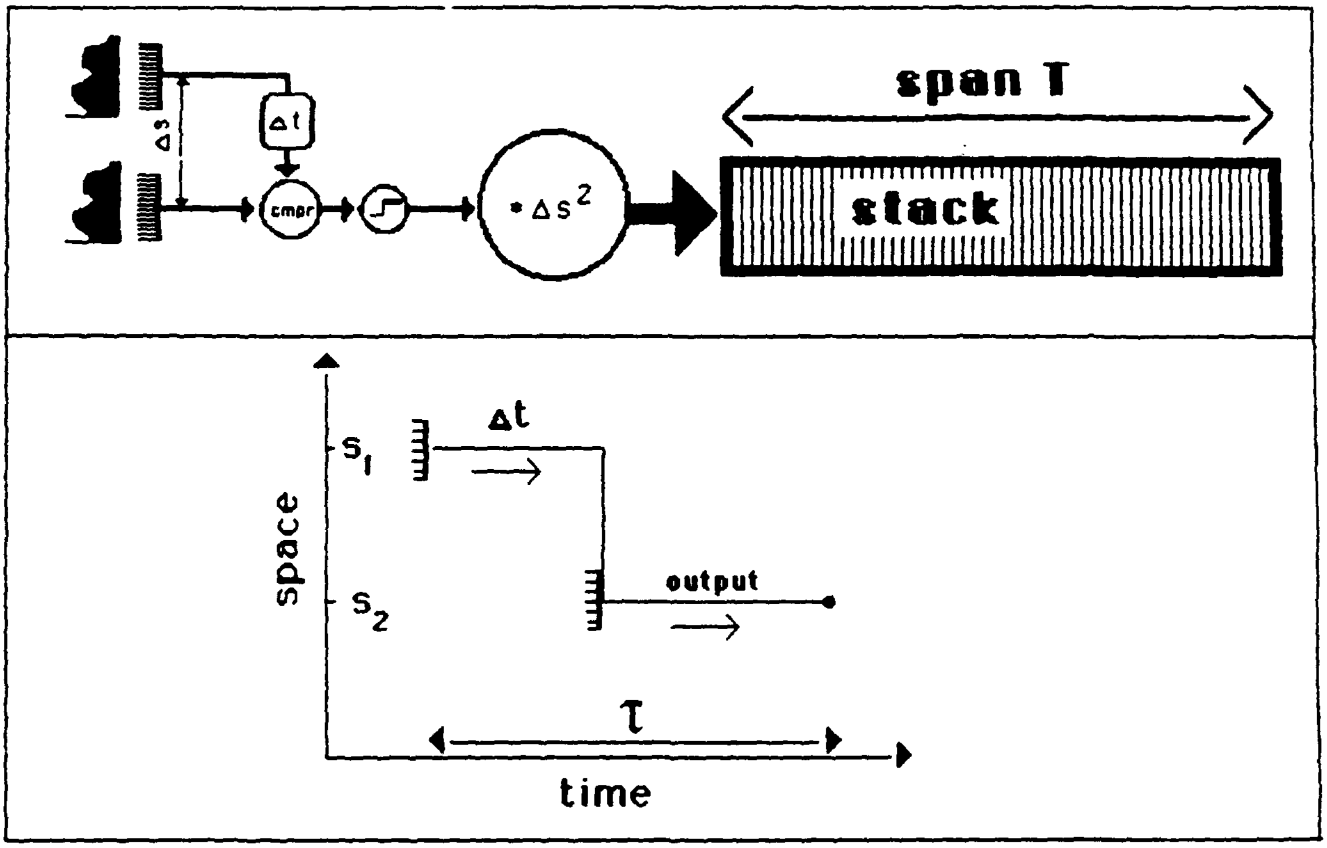


FIGURE 10

Project Seven:

Psychophysics and physiology of motion perception*

Robert Sekuler, Stuart Anstis,
Oliver J. Braddick, Thomas Brandt,
J. Anthony Movshon, and Guy Orban

*Chapter for L. Spillman and J.S. Werner (eds.)
Neurophysiological Foundations of
Visual Perception, Academic Press, 1989

I. Introduction

Of the many important contributions to the study of motion perception, three stand out as towering landmarks in a long and provocative scientific literature. Like their physical counterparts, these intellectual landmarks can serve to orient travelers who are unfamiliar with the surrounding territory. Also, these landmark contributions have influenced much of the current work on motion perception. As a result, they provide a convenient and natural entree to our discussion.

The first modern study of motion perception was done by Sigmund Exner a century ago. Although other scientists before him also made significant contributions, Exner (1888) was the first to appreciate that, as a perceptual quality, motion was special. Physically, motion can be described as a spatial change over time. But Exner demonstrated that perceptually, motion was not merely the stepchild of the perception of space and and the perception of time. In one study (1888), Exner placed two sources of electrical sparks so close to one another that an observer could not distinguish the two. Despite the impossibility of resolving the sources spatially, when the sparks were presented with the appropriate time interval, observers experienced compelling motion —the spark seemed to move from one location to another. In other words, observers experienced motion even though they could not spatially resolve the two endpoints of the motion. Exner also succeeded in demonstrating that observers could experience motion even though they could not temporally resolve its sources.

The second landmark in our brief history of motion perception is the studies Max Wertheimer reported in his 1912 monograph. Unfortunately, this monograph is more often quoted than read. But for anyone who delves into it, Wertheimer's monograph is a remarkable source of stimulation. Many readers will be familiar with the monograph's main experiment: a compelling sensation of motion can be produced by a brief, sequential presentation of first one then the other of two, spatially adjacent lines. But the monograph also anticipates issues that many vision researchers are working on today. For example, Wertheimer provided a good demonstration of motion inertia, a phenomenon that is central to perceptual theory which will be described below. He also provided a good demonstration of hysteresis, a form of "memory" that has proven to be of much theoretical importance (Williams, Phillips & Sekuler, 1986). But that is just a small, highly selective sample; Wertheimer's entire monograph is worth reading for enlightenment as well as for stimulating research ideas.

And finally, the third landmark exploration of motion perception is Werner Reichardt's studies in the late 1950's and early 1960's. These studies constitute a genuine paradigm shift, in Thomas Kuhn's sense (Kuhn, 1970). Reichardt's elegant mathematical model (1961) shifted the field from its prior status -- an enterprise geared to the uncoordinated collection of interesting facts -- to a field with a consensus about research methods and priorities. Reichardt's model, for the first time, stimulated people to think about how the visual system might extract motion information from the stimulus on the retina. His work, though now a quarter-century old, remains very much alive today. Most current models of motion extraction are elaborations on Reichardt's original model. Basically, the scheme assumes that the visual system compares the signals that arise, over time, from different photoreceptors. If some pattern travels across the retina, its effect on receptor R_a, at time t will be strongly correlated with its effect on another receptor, R_b, at some slightly later time, t+D. There will be a strong cross-correlation, with lag D, between the signals from the two receptors. Note that the spatial separation between the two receptors acts to delay one signal relative to the other. Figure 1 illustrates the kernel of Reichardt's model. Note that this simple scheme makes use of only two receptors (shaded rectangles in each panel). Poggio and Reichardt (1973), extending the basic scheme, have shown that a motion detection model with *n* inputs and a single output can be reduced to the sum of 2-input pairs, the case illustrated in Figure 1. (See Sekuler, Pantle and Levinson (1978) for more details.) Reichardt's work opened the way for the development of detailed, quantitative accounts of motion perception (Reichardt, 1987).

Figure 1 about here

Although the mathematical precision, clarity, and force of Reichardt's contribution makes today's work on motion more coherent, there is still an enormous variety of approaches to motion. There's a good reason for this variety, despite what seems to be quite a broad consensus.

The Many Functions of Motion

Since motion plays so many different perceptual roles, researchers can emphasize or concentrate on certain aspects only. Such choices necessarily lead researchers along different paths. Let us briefly review some of motion's many roles (see Nakayama, 1985, for a more thorough treatment).

Motion is particularly important for segregation of figure and ground. If an object moves relative to a background, producing differential speeds or different directions in the retinal image, the visual system converts those differences into perceptual separation of figure and ground (shape from motion; see Chapter 10 in this volume). Motion also segregates, or sorts, objects into different depth planes (depth from motion; again, see Chapter 10 in this volume). When any single region of the retina is stimulated by different velocities, the visual system is challenged (shearing; see Koenderink & van Doorn, 1978). Different velocities usually mean different objects --or parts of objects. But in the natural world, two (or more) objects cannot occupy the same place at the same time. The visual system seems to resolve the apparent contradiction of different velocities within a single region by assigning those different velocities to different depth planes. The processes responsible for such depth assignments, structure from motion, are currently being actively investigated by vision scientists, physiologists, and people interested in computer vision.

Finally, one of the main motives for studying motion perception is a desire to understand how motion helps us avoid colliding with objects, keeps us moving along the straight and narrow, and helps to maintain our posture. As will be shown later, motion's different functions probably require that a variety of different neural computations be carried out, most likely by different neural circuits.

A. Stimuli.

Scientific research is limited —or empowered— by the tools that are available. By probing it with a sufficiently complex and rich stimulus, the motion system can be forced to reveal its own richness. For this purpose, stimuli belonging to the family of random dot cinematograms are especially good. All members of this family have two features in common: first, a random spatial arrangement of their elements, which is designed to minimize visible contours; second, some rule or rules that govern the way in which those elements are displaced from one frame of the display to the next. [See Chang (1986) for a discussion of these stimuli.]

A pair of problems. Random dot cinematograms present a special challenge to the visual system in the form of the correspondence problem. The term "correspondence problem" denotes the challenge of matching elements in one frame with elements in a succeeding frame. From top to bottom, Figure 2a illustrates three successive frames of a random dot cinematogram. Some subset of dots from the first frame (top) has been shifted in the second frame (middle) and shifted again in the third frame (bottom). In Figure 2b the shifted dots are highlighted for ease of identification.

Figure 2 about here

Though effort is needed to find the subset of dots in Figure 2a, when the same frames are shown as a cinematogram the visual system extracts the subset with no effort at all. More specifically, if the three frames illustrated in Figure 2 were spatially superimposed and shown in rapid succession, the displaced set of dots would appear to move upwards and to the left. The visual system has little trouble extracting coherent motion of the shifted dots. Even with enormously dense and complex displays, perceived motion can still be easily extracted.

"Matching" may not be the proper term for what the visual system is actually doing, but it is a term in common use. There are various possible strategies for solving the correspondence problem. One strategy might be a point-by-point match. This approach may actually be employed when the cinematogram contains just a few elements. But when the cinematogram contains several thousand elements and only relatively few are being shifted in a coherent fashion, a point-by-point match becomes unfeasible. An alternative employs a more global strategy, one heuristic or another from the "bag of tricks" to which Ramachandran and Anstis (1986a) have called attention.

The correspondence problem is not the only challenge that the visual system must overcome in its quest to extract useful information about object motion. Another significant obstacle, the aperture problem (pp. xx), arises from the limited field of view, or receptive field, assigned to any visual neuron. Imagine that you see the world through the narrow local window of a single receptive field. This restricted field of view necessarily creates ambiguities. Suppose an infinitely long edge is moving through the receptive field. Any one of a large number of combinations of directions and speeds could mimic perfectly the velocity of that edge. So, to that neuron, many combinations of directions and speeds ought to be indistinguishable. Yet, except in some very special circumstances, perceivers do not make the sort of confusions that one neuron would. As will be seen later in this chapter, Movshon and his colleagues (1986) have developed an ingenious scheme to circumvent this apparent neuronal limitation.

The ease or difficulty with which one experiences motion in displays depends upon a number of spatial and temporal variables. For example, if the elements from one frame to another are shifted by very large steps, the sensation of motion breaks down –instead of motion, one set of dots seems to disappear and then reappear at a different location. This upper limit is now called d_{max}, the largest displacement between successive presentations for which observers still obtain a coherent sense of motion. The existence of such a limit has been known for some time; Wertheimer (1912) noted and Korte (1915) formalized it in one of his laws of apparent motion. More recently, this spatial limit, d_{max}, has has become an indispensible tool for understanding motion perception. Among other virtues, d_{max} can be quite useful for bridging the gap between psychophysics and physiology. As this chapter shows, measurements such as d_{max} can be made in several different domains: on single neurons, in human and animal observers. Such comparisons, allow connections to be made across the domains. At various points in the chapter, we will note particular linking propositions, statements that assert some link between physiological (ϕ) and psychophysical (Ψ) domains. (See Chapter 2, this volume.)

Returning to the nature of test stimuli, it is worth pointing out that random dot cinematograms can vary in a great many different ways. Consider first, the life time (exposure duration) of each individual element. In some random dot cinematograms, individual elements have a short life expectancy; they exist for a short time and then disappear to be replaced by other random elements (Mather & Moulden, 1980; Andersen & Siegel, 1988; see Chapter 3, this volume). This renewal scheme minimizes the probability that individual dot paths are being tracked. For example, one could not compute a dot's direction by comparing the points at which it entered and exited the display area.

The successive displacements of the elements in a cinematogram can be governed by various sorts of rules. For example, a simple rule can be used, causing all the moving elements to be displaced in unison —all in the same direction and at the same pace. This represents the extreme of narrow-band directional content: only one direction is present. However, there are other possibilities. Random dot cinematograms can contain many different, spatially intermingled, directions of motion, all present simultaneously. For example, the elements in a given local area may be displaced over a series of frames, not by some constant step and direction, but by directional values drawn from a distribution with some given mean direction and covering a range of directions (Watamaniuk, Sekuler & Williams, 1988).

With this latter scheme, a perceiver may experience two contradictory percepts: a)different directions of local motions, and b)a coherent flow in the direction of the mean of the directional distribution (Williams & Sekuler, 1984). One sees individual dots moving randomly, but at the same time one also perceives the overall flow of the dots in some dominant direction. This global percept enables one to use the concept of metamerism to study the mechanisms of motion perception (Richards, 1979). Two stimuli are said to be metameric if, despite physical differences, they are perceptually indistinguishable. Under appropriate circumstances, metamers reveal what information the visual system retains and what information it discards. Although best known and exploited in color vision, metamers also exist in the domain of motion: radically different distributions of directions are able to produce perceptually indistinguishable motions.

To use psychophysical "confusions" to study the number or type of underlying mechanisms, one postulates a particular type of linking proposition, the Converse Identity proposition (Teller, Chapter 2, this volume). One formulation of the proposition states that "statistically indiscriminable sensations imply statistically identical states of the nervous system." Symbolically, this proposition can be stated as

$$\text{Identical } \phi \text{ -> Identical } \Psi.$$

An Important Spatial Limit to Motion

In analyzing human motion perception and its possible neurophysiological basis, we have learned a lot from experiments on apparent motion using random dot cinematograms such as the one illustrated in Figure 3. As indicated earlier, a key parameter from these experiments is d_{max}, a spatial limit on apparent motion. This parameter has particular significance when one tries to relate the psychophysics of motion to the physiology of direction-selective neurons; it presumably represents the spatial range of the interactions that underlie directional selectivity within the receptive fields.

Figure 3 about here

Earlier studies (Braddick, 1974) suggested that d_{max} was a quite closely defined parameter. It seemed to fall between 15 and 20 minutes of arc, regardless of variations in the size and spacing of the elements of the random dot patterns. This measure is striking because it represents a distance much smaller than the range for apparent motion in the classic work of Wertheimer (1912). Recall that in those studies, Wertheimer, and others after him (e.g., Neuhaus, 1930), quantified the range for which one could see motion when spatially-offset lines or spots were alternated, and typically found this range to be at least several degrees of arc (see also Jung & Spillmann, 1970). The term "short range process" was coined by Braddick (1974) to indicate a motion process that would be particularly responsive to cinematogram dsplays. It was meant to contrast with the longer range process that yields perceived motion in patterns which contain a small number of clearly defined elements. Although recent work (see below) shows that the spatial limit cannot be thought of as an invariant 15-20 min arc, the idea of a distinct short-range process seems still to be valid.

In fact, two demonstrations show that, under the right conditions, d_{max} can be extended a good deal beyond 15-20 min arc. In one study, Baker and Braddick (1985) constructed random dot cinematograms in which dot displacements occurred within a pair of strips on either side of the fixation point. By varying the separation of the strips, d_{max} could be measured at several different eccentricities of viewing. The results in Figure 4 show that d_{max} increases with retinal eccentricity in an orderly fashion, and at 10 degrees from the fixation point, displacements as large as 90 minutes can be perceived as coherent motion. This increase is not entirely unexpected, since most spatial parameters of vision increase with eccentricity (see Chapter 9, this volume). Therefore the maximum range of motion detection is no exception to the rule. However, two features of this result are worth noting. First, the increase in d_{max} with eccentricity implies that direction of large displacements (or high velocities) can be perceived more accurately in peripheral than in foveal vision. This is one of the few ways in which the performance of peripheral vision is actually superior to that of the fovea. Second, the function relating d_{max} to eccentricity does not have the same form as that found for parameters such as minimum angle resolvable (acuity) [see Chapter 9, this volume]. Most likely, task-dependent variations in scaling with eccentricity suggest that different visual functions depend upon different subpopulations of visual neurons.

Figure 4 about here

In concluson, d_{max} should not be thought of as a constant; its value depends on the location tested in the visual field. Eccentricity is not the only variable that alters d_{max}; an equally potent variable is the pattern's spatial frequency content. The elements in the random dot cinematogram illustrated in Figure 3 had quite sharp edges, and this is important in determining the measured d_{max}. For example, if one takes such a cinematogram and adjusts the displacement so that it just exceeds d_{max}, the sensation of motion will cease, as one would expect. However, with the displacement still set at the same value, one can immediately restore the sensation of motion by simply squinting and thereby blurring the image. Thus, blurring, which removes the high spatial frequencies in the pattern, has effectively increased d_{max}. This may also explain why d_{max} is larger in the periphery of the visual field.

Cleary and Braddick (1985) tested the effects of spatial frequency on d_{max} in a more systematic way. They filtered random dot patterns so that they contained only certain narrow bands of spatial frequencies. In three different cinematograms, the center frequencies of these narrow bands were 1.3, 2.7 and 5.3 cycles per degree. Results with these three filtered cinematograms are shown in Figure 5. The y-axis shows the percentage of errors in observers' directional reports. The values on the x-axis are not expressed in terms of minutes of arc, as the values in some preceding figures had been. Instead, x- axis units are the number of cycles of the center frequency of the band, for each cinematogram. Plotted in this way, the three functions are virtually identical, and in particular d_{max} (taken as the lowest displacement for which error rate rises to 20%) falls at about the same value for all three cinematograms. However, this constant number of cycles will occupy very different spatial extents (e.g., one cycle of 5.3 cycle/deg cinematogram covers only about one-fourth the distance covered by one cycle of the 1.3 cycle/deg cinematogram). That is, the similar d_{max} values in Figure 5 imply a factor of four variation in d_{max} expressed in the usual, angular distance units. Chang and Julesz (1983) have reported rather similar results. Thus, within a given region of the visual field d_{max} is a function of the spatial frequencies present in the image. Roughly speaking, if a random dot cinematogram consists of big blurry patches, one can see them move over large displacements. There is a great deal of evidence that the visual system contains receptive fields, or channels, with different spatial frequency properties. These results suggest that at any particular location in the visual field, the overall scale of those receptive fields differ, not only in the scale of patterns to which they are sensitive, but also proportionally, in the scale of displacements that they can detect (compare the Chapter 9, this volume).

Figure 5 about here

There are some other significant features in these data. Figure 5 shows that in terms of the characteristic spatial frequency of a narrow-band cinematogram, d_{max} turns out to be just about one cycle of that frequency. (There are also some interesting oscillations in performance for displacements above one cycle, but we are not concerned with those here.) Several current theoretical models of motion processing imply that the limiting displacement ought to be somewhere between a quarter and half a cycle (e.g., Adelson & Bergen, 1985; van Santen & Sperling, 1985). The form of the data casts some doubt on such 'quadrature phase' models. It implies that correlation mechanisms that underlie motion perception are not necessarily confused by the similarity between one cycle and the next. Therefore, they must be using additional information than simply matching the locations of individual zero-crossings or peaks in the one-dimensional signal. This additional information might be contained in the detailed shape of the waveform, combined information from a range of orientations, or an extended area of the pattern.

A second important implication comes from considering the original, unfiltered, cinematogram. This contains a broad band of spatial frequencies, including the low frequencies which are known to yield a large d_{max}. However, d_{max} for this broad-band pattern can still be increased by blurring, which does not add to these low frequencies but simply attenuates the high spatial frequencies. Thus, we conclude that the presence of high frequencies (fine detail) can interfere with the use of motion information potentially available in the low frequencies. That is, although each spatial frequency channel has its own d_{max}, in this situation they do not act independently. This inaccessibility of information carried in the low spatial frequencies when high frequencies are present is reminiscent of the way a static picture can be made unrecognizable by segmenting it into sharp-edged blocks (with high spatial frequencies), as in the well-known 'Abraham Lincoln' demonstration (Harmon, 1971). When optically blurred, Lincoln's photo becomes immediately visible. Clearly, independent frequency channels are by no means the whole story, neither with respect to suprathreshold pattern recognition nor to motion perception.

What do these results imply for defining a distinctive short range process in motion perception? As we have seen, d_{max} varies as a function of the spatial frequencies in the retinal image, and as a function of retinal eccentricity. If this limit is so variable, does the division between 'short-range' and 'long- range' have any meaning at all? Perhaps the perceived motion ascribed to a distinct long-range process occurred in conditions when low spatial frequency information could be used, and consequently d_{max} was large. However, the short range process was not defined solely in terms of its spatial limit. Even though various manipulations can affect the value of d_{max}, certain temporal properties seem to be characteristic of the short range process.

Take for instance the eccentricity variation mentioned before. Figure 6 shows the results of measuring d_{max} at different eccentricities while also varying the interval between the first and second exposure. As the interval approaches 100 msec all the curves fall off in a similar manner. This temporal variation does not seem to change with eccentricity. And indeed, there is a similar effect with variation in spatial frequency: similar limiting intervals seem to hold regardless of the cinematogram's spatial frequency content. Again, with respect to the more classical kind of apparent motion displays (e.g., Wertheimer, 1912) these are relatively short intervals. Therefore the distinctive temporal aspect of the short-range process holds up across spatial variations, and a different, 'long-range' process must be invoked to account for the apparent motion with single elements that can be seen with considerably longer delays.

Figure 6 about here

At the beginning of this section, it was suggested that d_{max}, a psychophysical variable, should be related to physiological measures of the range of direction-specific interactions within receptive fields. Figure 7 illustrates some data which may allow us to make this connection. The basic strategy is to measure a neuron's response to a stimulus that is displaced laterally from one brief presentation to the next. As the displacement gradually increases, one notes the displacement at which the responses cease to be directionally selective (e.g. the response to a displacement to the left becomes as strong --or as weak-- as that to a displacement to the right). This procedure is analogous to the psychophysical assessment d_{max}. The data shown in Figure 7 were gathered by Mikami, Newsome and Wurtz (1986) from single neurons in macaque monkeys. The stimuli were not random dot patterns as used in the psychophysical d_{max} experiments, but thin bars that appeared to step across the visual field in a series of flash exposures. Triangles represent measurements on directional cells in Area V1, known as the primary visual cortex. As the regression line (dashed) indicates, the physiological analogue for d_{max} varies somewhat with retinal eccentricity. Mikami, Newsome and Wurtz also studied neurons in cortical area MT, which is believed to be specialized for motion processing. In Area MT, the analogue for d_{max} not only grows more rapidly with eccentricity, but also reaches higher values than for V1 cells at comparable eccentricities. Mikami, Newsome and Wurtz (1986) have plotted some of Baker and Braddick's (1985) data on these same axes (the squares). The human psychophysical data seem to correspond much more closely to the maximum displacement for cells in Area MT than in Area V1.

Figure 7 about here

Obviously there are problems in relating the performance of single cells to an observer's performance on some psychophysical task. For one thing, the observer presumably uses the signals coming from a very large number of cells. There are usually differences between stimuli. For instance, the stimulus used by Mikami and his associates made many steps as it traversed the receptive field. As we shall see below, psychophysical d_{max} is higher when the stimulus takes more than two steps. Nonetheless, it is striking that psychophysically we can detect the direction of displacements that are greater than those that elicit directional responses from any cell in Area V1. It is equally striking that psychophysical performance falls within the range of displacements processed by Area MT cells. Apparently, a major portion of the "machinery" for extracting motion information from successive exposures lies beyond Area V1, that is, beyond the first major direction-selective stage in the magnocellular stream (see Chapter 5 in this Volume).

When directionally-selective units are probed, their directional selectivity does not depend on the stimulus passing right across the receptive field. Rather, any small region of the field shows directional responses (see Barlow & Levick, 1965). In fact, studies such as those by Mikami et al. in macaque cortex (1986), show that the largest displacement that produces a directional response is normally considerably smaller than the cell's receptive field. That has led to the idea that a receptive field is made up of many local subunits each of which is directionally selective (Emerson, Citron, Vaughn & Klein, 1987). Each subunit can generate a signal that contributes to the cell's overall directional response. Yuille and Grzywacz (1988) offered an explicit computational account of how motion information might be integrated over space. Their model gives an excellent account of various psychophysical demonstrations that neighboring regions in the visual field interact cooperatively to produce an overall sensation of motion (Chang & Julesz, 1984; Williams, Phillips & Sekuler, 1986).

Under certain conditions, stochastic displays can give rise to a percept of global motion, an effect that can shed considerable light on the visual system's strategy for integrating motion over space. To examine this phenomenon, Williams and Sekuler (1984) developed special random dot cinematograms in which all dots drew their successive displacements from a rectangular distribution characterized by some particular directional range. On any frame of the display, the direction in which any single dot moved was i)independent of its own history of movements, and ii)independent of the movements of other dots. When the distribution of directions covered a broad range of directions –for example, 270-360 deg– the observer saw only the local, random motions of individual dots. When the distribution of directions covered a narrower range –for example, 90-180 deg– the observer continued to see those local random motions, but now also saw a global, coherent flow in the general direction of the mean of the distribution.

Williams and Sekuler (1984) measured the probability with which global motion was seen, as a function of the directional distribution s range. They found that as the distribution changed, so did the percept –from random, incoherent motions to global flow, or vice versa. Moreover, the perceptual change tended to be quite abrupt, making for frequency-of-seeing curves that were quite steep.

After exploring a number of the parameters that determined when stochastic motions would yield a global percept of flow, the obvious challenge was to characterize the system that might produce such behavior. Williams, Phillips and Sekuler (1986) suspected that cooperativity might be involved. Since hysteresis, a form of memory, is regarded as a reliable marker for cooperativity, they set out to determine whether the percept of global flow exhibited hysteresis.

In their typical experiment, a trial began with a random dot cinematogram that contained either a narrow- or a broad-distribution of directions. Then, after some random interval, the distribution changed slowly over successive frames. This change in the direction distribution caused the percept to change, either from global flow to random noise (when the starting distribution was narrow), or from random noise to global flow (when the starting distribution was broad). The dependent measure was the direction-distribution at which the percept changed from one state to the other.

The basic finding is simple: the dependent measure varied strongly with the starting state of the stimulus (and, correlatively, the starting state of the percept). If the initial conditions had promoted a percept of global flow, the percept switched states at a relatively broad distribution; if the initial conditions had not promoted a percept of global flow, the percept switched state at a far narrower distribution. Quantitative estimates of the effect of starting state were obtained under several different stimulus conditions. After studying several control conditions, Williams, Phillips and Sekuler (1986) concluded that the percept, once established, did indeed exhibit a resistance to change –that is, the percept exhibited hysteresis. The results of these experiments were well fit by a model that involved cooperative and competitive interactions among direction-selective units. A network comprising such interactions decides among alternative percepts on a "winner-take-all basis" (Feldman & Ballard, 1982).

One should expect that the visual system would combine motion information not just over space but over time as well. In integration over space, there is one range – within the subunit of the receptive field – that is related to d_{max}, but there is also a larger range over which there are interactions among separate subunits. Analogously, there might be two time constants in the motion system. One time constant might relate to the maximum interval for a single displacement. However, if the subject is presented with a sequence of more than two exposures, information may be integrated over a period much longer than the interval between a single pair of display frames. Figure 8 illustrates this idea with some data from Snowden and Braddick (1987). Using random dot cinematograms, d_{max} was measured as a function of the number of successive exposures per trial. Note that as this number increases, d_{max} increases as well. As Figure 8 shows, d_{max} increases up to between four and six displacements (see also Nakayama & Silverman, 1985). Clearly, the motion system is gaining extra information from temporal integration over successive displacements. The detector's basic directional response requires two exposures within less than 100 msec, but Figure 8 shows that this response is enhanced by temporal integration over at least 300 msec.

The results of Figure 8 show integration over time, but one should not conclude prematurely that the asymptotic performance is determined by the temporal limits of integration. At least over the range shown in Figure 8, when the entire train of displacements is speeded up or slowed down, d_{max}'s asymptotic number of steps remains constant. The time to reach asymptote varies between about 100 and 400 msec, depending on the rate of presentations. Of course, motion does inevitably involve both time and space: in taking n steps, each of a given d_{max}, the stimulus traverses a particular distance.

Figure 8 about here

Figure 9 is taken from Mikami, Newsome and Wurtz's (1986) experiment on the relation between receptive field width and d_{max} for macaque Area MT neurons. As mentioned before, the receptive field widths tend to be much larger than d_{max}, implying some form of subunit structure. The graph plots the ratio between field width and d_{max}. Note that this ratio shows a shallow gradient with eccentricity; between four and seven steps of d_{max} would fit within one receptive field. This finding resembles the kind of asymptotic value found by Snowden and Braddick (1987), suggesting that the limit might be set by the width of the receptive field.

Figure 9 about here

However, further experiments by Snowden (19xx) suggest that the limiting factor may be neither spatial nor temporal, but simply a constant number of steps. Such a limit might reflect the effectiveness with which signals can be propagated from one detector to the next, across a cooperative network. The idea that linked detectors combine information by means of mutual facilitatory and/or inhibitory interactions can be contrasted with the simpler idea of subunits, each having an independent directional response, whose outputs are summated in the motion detector. (Of course, such combinations must have limits on its temporal and spatial range, even if they are not the major factor in these experiments.)

These results imply that in order to understand how neural structures determine psychophysical performance in motion perception tasks, it will not be sufficient to examine the performance of an isolated motion detector. Our perception of motion depends on integrating a number of local neural responses. Clearly, further psychophysical experiments need to be done to characterize this integration. For instance, how far can successive steps at different effective velocities, or in different directions, be integrated? Hopefully, there will also be advances in physiological knowledge that will clarify the neural basis of this integration and of other factors that affect d_{max}.

A. Correspondence challenges and correspondence solutions.

In order to extract apparent motion from complex displays, the visual system somehow solves a very difficult problem. Among the thousands of possible element-to-element matches, only one is correct. How does the visual system determine which parts of successive images reflect a single object in motion?

Ramachandran and Anstis (1986a) have suggested that early stages of visual processing of motion uses various heuristics, or rules of thumb, that the human visual system has acquired through millions of years of evolution. These heuristics have been adopted not for mathematical elegance or aesthetic appeal, but merely because they worked. One can learn much about these rules of thumb by watching the visual system as it struggles to solve the correspondence problem.

These rules reflect the fact that in the real world objects move in characteristic, predictable ways. For example, if one's arm moves, neighboring parts of the arm tend to move together. Or, as a football spirals through the air, its parts tend to travel en masse (no piece of the pigskin or the laces are likely to peel off on its own independent course).

The visual system seems to make the correct assumption. At a macroscopic level the physical world is not a chaotic, amorphous mess. The visual system capitalizes on the world's predictable physical properties and limits the matches it must consider by dealing only with matches that would yield perceptions of motion that are plausible in the real, three-dimensional world. Later on, we will return to speculate how this scheme might be implemented. To examine the notion that the visual system assumes the world has order, Anstis and colleagues fashioned motion displays that could be interpreted in more than one way and then observed how this ambiguity was resolved (Anstis and Ramachandran, 1986a; Ramachandran and Anstis, 1986b). The resulting percepts –or interpretations– suggested that the visual system was making three different but quite sensible assumptions about the real world: 1) inertia – that moving objects tend to continue moving in the same direction, showing minimal changes in velocity over time; 2) rigidity – that extended surfaces tend to move all in one piece, showing minimal changes in velocity over space; and 3) that moving objects tend to cover and uncover predictable regions of the background.

Assumption One: Inertia. The visual system makes one assumption that may remind you of Newton's first law of motion: objects in motion tend to continue their motion along the same path. (Note the resemblance between this statement and the Gestalt law of good continuation [Bruce & Green, 1985]). Perceptually, once motion is experienced there is a tendency to continue to experience it, even after the motion has actually stopped. Wertheimer's monograph (1912) offered an intriguing demonstration of this fact. He produced apparent motion by a series of alternating presentations of two vertical lines. During the alternations, without warning to the observer, Wertheimer occluded one, leaving the remaining line to appear at its normal time in the sequence. Observers continued to see motion for several "cycles" after the line had been occluded. Anstis and Ramachandran (1986b) gave an elegant demonstration of this phenomenon in the case of rotary inertia. Figure 10 illustrates one arrangement that shows inertia in rotary motion. First, consider control measurements, that evaluate motion perception with no inertia. The display alternates between a pair of crosses that are rotated 45 degrees relative to one another. These crosses can be thought of as a plus sign and a letter x. When these two spatially overlapping figures are alternated, the apparent motion they set up is ambiguous. Rotation is seen, but it can be either clockwise or counterclockwise (anti-clockwise, if the demonstration is performed in Great Britain). The bottom part of the figure illustrates what happens when an inertia-inducing constraint is added. This constraint is another, tilted, cross, oriented toward eleven o'clock on the watch face. This constraining cross, in frame 1, precedes the sequence of the other two crosses. Now the first two presentations, a tilted cross followed by a plus sign, produce strong motion in a clockwise direction. Interestingly, this strong apparent motion continues when the third element, the letter x, is presented. The first, unambiguous jump, imparts a perceptual rotary inertia that converts the previously ambiguous motion into one that inevitably is seen as clockwise. This may be termed a form of motion bias (priming).

Figure 10 about here

Assumption Two: Rigidity. Another assumption that could limit possible correspondences is the assumption that objects are rigid; that is, all points on a moving object are assumed to move in synchrony. Though many interesting objects are not rigid in the strict sense, most exhibit at least some local rigidity —a tight coupling between the movements of closely neighboring parts. The tendency for neighboring components to move in similar ways lends considerable redundancy to the motions of neighboring elements within image space. This redundancy would make it economical for the visual system to extract salient features, such as clusters of elements, rather than individual elements, from a complex display and then search for corresponding features in successive images. This strategy, if it could be implemented, would certainly reduce the number of potential matches without increasing perceptual errors. Take a leopard leaping from a branch of one tree to a branch of another. (Note the resemblance between this statement and the Gestalt law of common fate [Bruce & Green, 1985].) According to the rigidity assumption, a viewer who picks out any salient feature of the leopard, such as its basic shape, and finds the same feature in a second frame need not compare every black spot on the animal at moment t_0 with every single black spot at moment t_1. A highly efficient system might take advantage of the spatial redundancy by attempting to match features on a coarse scale. Consistent with this idea, Ramachandran, Ginsburg and Anstis (1983) found that the visual system often detects correspondence between regions of similar low spatial frequencies before it detects more detailed outlines or sharp edges. The same heuristic might also account for the way in which a cinematogram's spatial frequency content influences d_{max} (see above).

Assumption Three: Covering and Uncovering. The visual system appears to make a third assumption, which is a corollary of the other two: a moving object will progressively cover and uncover portions of a background. J.J. Gibson (1966), among others, has called attention to the importance of this fact. When an object, which is normally opaque, temporarily occludes a background, the background still exists; it does not disappear. To see how the third assumption affects perception, consider Figure 11. The left panel of the figure illustrates a display in which a triangle and a square below it are presented and then are replaced by another square adjacent to the triangle and directly to its right. As the right panel suggests, one sees the triangle appear to move horizontally and to hide behind the obliquely moving square, which now appears to occlude a triangle that is not, in fact, being displayed. The visual system seems to assume that an object continues to exist, even if the system has to fabricate the supporting evidence (Anstis and Ramachandran, 1985).

Figure 11 about here

But consider even more complex stimuli. What strategy could the system adopt when presented with many objects simultaneously in apparent motion? The visual system behaves economically, perceiving all objects in a field as moving in the same direction, unless there are unambiguous cues to the contrary.

Figure 12 provides an example of another spatial constraint, one that operates on a more global scale (Ramachandran and Anstis, 1985). The figure shows nine ambiguous quartets of dots, with two dots (either black or shaded) from each quartet appearing in each frame. Under proper conditions, observers report that the dots in each pair sometimes move vertically and sometimes horizontally, though in opposite directions. The percept fluctuates more or less randomly. The interesting point is that all the quartets move in the same direction at any given time. If the dots in any one quartet appear to move vertically, the dots in all the quartets do likewise. Then suddenly they all change step together and move horizontally; the dot quartets entrain each other. There is a strong tendency towards seeing spatial coherence, or if you like, uniformity across the field (see also Chapter 10, this volume).

Figure 12 about here

The visual system behaves as though it took advantage of certain rules of thumb about the properties of objects in the real world. Naturally, if this viewpoint is more than an interesting metaphor, we need some idea of how such behavior is possible. How might such "assumptions" be implemented, either in neural hardware or in software? Figure 12 gives some hint of a reductionistic explanation of this phenomenon. As was noted earlier, Braddick has shown that when the total excursion of some stimulus occurs in a series of small, successive displacements, d_{max} increases. The result is that one is more likely, than otherwise, to see motion in a straight line. Among the interesting questions that remain, though, none is more intriguing than the question of the genesis of these assumptions. Are they represented in neurons that are "hard-wired" from birth to implement those assumptions or strategies? Or do those assumptions become wired into the system as the result of some kind of natural selection at the neural level, a kind of neural Darwinism (Edelman, 1987; see Chapter 12, this volume).

Because this process depends upon the viewer's own experiences with his or her environment, the resultant neural connections would be likely to reflect the properties of objects and motions in that environment. Alternatively, does the perception of motion require some higher level of cognition? Time --and further research-- will tell.

Cortical Mechanisms

A. Initial stages.

While various psychophysical phenomena of motion perception are still fresh in mind, let us consider some of the neuronal mechanisms that might contribute to the psychophysical effects that have been discussed so far. Basically, in the visual cortex of cat and monkey three types of cells seem most likely to play major roles in motion perception. Two of these cell types -- direction-selective and velocity-tuned-- seem well-suited to provide estimates of local motion as opposed to global motion. The third, more recently discovered type, is a motion segregation or parallax cell. This type of cell may be especially relevant to several of the psychophysical effects described earlier.

When one asks what sort of cells contribute to perceived motion, the first answer that comes to mind is the direction-selective cell (Pasternak, 1986, 1987; Pasternak & Leinen, 1986). In fact, a preponderance of direction-selective cells in one region, such as Area MT, suggests that the region is involved in processing motion information (Albright, 1984; Newsome, Wurtz, Dursteler & Mikami, 1985; Newsome & Pare, 1988). But directional selectivity is not easily defined. It is not only a matter of having a preferred direction of motion, a direction to which the cell responds more vigorously than to any other. Nor is it only a matter of directional asymmetry, in which the cell responds strongly to one direction and little or not at all to the opposite direction. Rather, directional asymmetry may be characterized by some ratio of responses to the motion in the optimal direction over responses to motion in the opposite direction. One commonly used formula is

$$\text{Direction Index} = (R_{pd} - R_{npd})/R_{pd} * 100,$$

where R_{pd} is the net response to stimulation in the preferred direction, R_{npd} is the net response to stimulation in the direction opposite the preferred direction, and "net response" signifies the difference between the response elicited by the stimulus and the mean spontaneous activity of the cell.

By such an index, cells vary widely in directional asymmetry. Various researchers have advocated using a high value of this ratio, typically an index of 50, as a cutoff bewtween direction selective from nonselective cells. However, cells' indices of direction asymmetry form a continuous distribution, not a bimodal one.

To make matters even more complex, for many cells, direction selectivity depends on the speed of the moving test target (Orban, Kennedy, & Maes 1981). One can then argue that statements about any cell's directional selectivity must be contingent statements, specifying the velocity for which selectivity has been assessed. For this reason, Orban et al. (1981) introduced the mean direction index (MDI) which is a weighted average of direction indices measured at different velocities, using the response strength at different speeds as weighting factors. Finally, the direction selectivity of some cells changes with the the sign of the target's luminance contrast (Albus, 1980; Yamane, Maske & Bishop, 1985; Orban, Gulyas, Spileers & Maes, 1987). For such cells, a bright bar on a dark background may yield a different index of selectivity than will a dark bar on a bright background. Perhaps the mean of the mean direction indices for light and dark bars might be a good index of the overall direction selectivity of a cortical cell.

The properties of direction selective cells might help account for some of the phenomena of apparent motion. A cortical cell's direction selectivity depends on interactions between distinct regions within the receptive field. Indeed, in Area 17 of the cat, if one masks the entire receptive field except for a central strip some 0.3° wide, the mean direction index of cortical cells is reduced to that of LGN cells. If a moving target is illuminated stroboscopically, direction selectivity to that target is abolished if the gaps between successive flashes are separated too much --either in space or in time (Duysens et al., 1988). This may explain why the short range motion process in apparent motion operates over short spatial and temporal intervals.

It is important to appreciate that not all direction selective cells are actually involved in encoding motion of an outside object. Indeed most physiological studies have used a single, isolated stimulus to measure direction selectivity, an artificial condition quite different from those occurring outside the laboratory. Recent results (Orban, Gulyas, & Vogels, 1987; Orban, Gulyas, & Spileers, 1988) demonstrate that in about half the cells in Areas 17 and 18 of the cat, direction selectivity for a foreground stimulus is modified dramatically by the motion of a textured background stimulus (see the solid curve in Figure 13A and the dashed curve in Figure 13C). These cells have been implicated either in motion segregation (see below) or in the extraction of depth from motion (Orban, Gulyas & Vogels, 1987). (Related observations have been made by von Gr nau and Frost (1983) in cat lateral suprasylvian gyrus and by Allman et al.(1985) in Area MT of the owl monkey.) Other cells, for which the direction selectivity does not depend on background motion, are most likely to encode motion of an object in the world. Presumably, they could signal the direction of object motion per se, without being strongly affected by the particular moving background against which the object happened to appear.

Figure 13 about here

Another type of cell that may be important in motion perception is the velocity tuned cell (Figure 14). Note that here the cell's response is a distinctly non-monotonic function of velocity, for both light and dark moving bars. Velocity-tuned cells typically respond optimally to some intermediate, moderate level of stimulus velocity. Orban and colleagues (Orban et al., 1981; Duysens et al., 1982) showed that at very high or very low velocities, velocity tuning was absent in Areas 17, 18, or 19 (Figure 15). The same holds for Area MT of the monkey where many cells are velocity tuned (Maunsell & Van Essen, 1983a,b,c).

Figure 14 about here

This observation leads one to expect that velocity discrimination, measured psychophysically, might also be best at corresponding, moderate stimulus velocities. Velocity discrimination measured in humans, cats and monkeys confirm that this is so (Orban et al., 1984; Vandenbussche et al., 1986a, 1986b).

In each species, the just noticeable difference in velocity is a U-shaped function of reference velocity. Also note that whereas humans and monkeys give very similar results (minimal just noticeable differences of about 5-7%), cats do more poorly overall (minimal just noticeable differences of about 50%), although under optimal spatiotemporal conditions, they may discriminate differences in velocity of about 15% (Pasternak, 1987). It is worth noting that these U-shaped functions remain invariant when measured with random dot patterns rather than moving bars (DeBruyn & Orban, 1986).

Striking though the analogy is, could the resemblance between psychophysical data and physiological data be merely coincidental? Of course one can never rule out such a possibility. However, one can strengthen the argument by exploring other dimensions of analogy. In particular, one can exploit the fact that velocity tuned cells do not constitute an entirely homogeneous class. Orban (1985), working in both monkey and cat, has shown that the optimal stimulus velocity for velocity-tuned cells varies with receptive field eccentricity. Figure 15 groups cells into three different ranges of eccentricity, 0-5 deg, 5-15 deg, and greater than 15 deg. Note that the optimal velocity increases with retinal eccentricity. One would expect velocity discrimination to show a similar dependence on eccentricity. Orban et al. (1985) found, with human observers, that indeed this is the case.

Figure 15 about here

To push the analogy even further, note that velocity-tuned cells lose their tuning when they are tested with slow (2-15 Hz) stroboscopic motion (Figure 16). Human observers show a comparable disruption of velocity discrimination when they, too, are tested with stroboscopic motion (Figure 17). A striking example of this is provided by MacKay's (1958) displacement illusion.

Figures 16 and 17 about here

In contrast to direction- and velocity-tuned cells there is a third group of cells that ordinarily shows no selectivity for direction, but does show selectivity under special conditions. Recognizing their potential perceptual role, Orban and Gulyas (1988) have called such cells "motion segregation" cells. Although motion segregation cells fall into several distinct classes (Orban, Gulyas, & Vogels, 1987), only one will be presented here, the so-called "anti-phase" cell.

When anti-phase cells are tested with a moving bar in the conventional manner --a single moving stimulus with no background movement-- they show no direction selectivity. At first glance, then, one might falsely think that these cells play no role in motion perception. However, the cell's response does change markedly when a moving background is introduced. In fact, in the presence of a background of moving, random noise, these cells become strongly directionally selective and this selectivity is quite complex (Hammond & Smith, 1982, 1984; Hammond, Ahmed, & Smith, 1986).

Whatever the direction of the moving background, and whatever the direction of a moving bar superimposed on that background, the cells respond most strongly when the bar moves in a direction opposite to the background motion (Figure 18). Because they are selective for target motion in a direction opposite to background motion, these cells can be labelled "anti-phase cells" (Orban, Gulyas & Vogels, 1987).

Figure 18 about here

Anti-phase cells fall into two classes. In Area 17 of the cat (Figure 18), and in Area V1 of the monkey, anti-phase cells exhibit selectivity only along one particular axis of movement. The other class of anti-phase cell, found in Area V2 of the monkey, exhibits selectivity regardless of the axis of motion. So long as foreground and background motions occur in opposite directions, it does not matter much what either direction is. This second class of anti-phase cells resembles the opposed-motion cells found by Frost and Nakayama in the pigeon tectum (1983).

What perceptual role might be played by motion segregation cells, including anti-phase cells? They all share a potential for signalling the presence of a difference, in speed or direction, between target and background motion. Commonly, such differences arise when target and background lie in different depth planes. Moving objects occupying different apparent depth planes and travelling in different directions set up shearing patterns of optical flow (Nakayama, 1985). Under some conditions, such shearing gives rise to strong kinetic contours, separating motion in one depth plane from motion in another. The perceptual conditions needed to see these kinetic contours have been studied by Koenderink and van Doorn (1978), among others. Those studies add support to the idea that direction segregation cells, rather than direction-selective or velocity-tuned cells, are involved in the creation of kinetic contours (Orban & Gulyas, 1988). For instance, if one measures the difference in direction of travel for two adjacent random dot cinematograms necessary for producing a kinetic contour, one finds that the direction difference has to approach 30 degrees, a value some 20 times higher than the difference threshold for direction. Interestingly, such large critical differences in direction are precisely what one would expect if motion-segregation cells, not direction-selective cells, played a key role in kinetic contours. Thirty degrees is about the smallest difference between foreground and background motions that elicits a strong response from motion-segregation cells.

The lesson, then, is that motion involves a great many different features, and that the nervous system makes use of several different cell classes to produce those features. Direction selectivity, though surely important, is not the alpha and omega of motion perception.

B. Area MT and the aperture problem.

One of the major problems of motion analysis is the way in which local motion signals are integrated to provide information about the motion of complex objects ar ' patterns. This class of problem was termed the "aperture problem" because it is readily made explicit when considering measurements of motion made through finite apertures (Movshon, Adelson, Gizzi and Newsome, 1986). Figure 19a illustrates the problem by considering the motion of two diamond figures, one moving down and one moving to the right.

Figure 19 about here

Although the global motion of these two figures is quite different, a local measurement of motion made in the circles drawn on the lower right-hand border of each diamond would yield the <u>same</u> value in each case. The local motion of a border is usually seen as being orthogonal to the border, as shown by the arrows linked to the circular apertures of measurement. This situation is formalized in the lower diagram of Figure 19a, a graph in which the angle of a vector represents direction of motion and its length represents speed. The local measurement of motion made in each of the apertures is not sufficient to define the motion of the whole object: there is an <u>ambiguity</u> concerning the motion measured locally. The true motion of the border consists of the measurable component <u>orthogonal</u> to the border and some unmeasurable (and therefore locally unknown) component <u>parallel</u> to the border (dashed line in Fig. 19). The measurable component is represented on the figure by the oblique vector directed down and to the right, and the unmeasurable component is represented by the dashed line orthogonal to it. The motion of the local border thus does not specify object motion completely, but imposes the constraint that the motion of the object containing the border must fall somewhere along the dashed line. The true motions in the two cases illustrated (the vertical and horizontal vectors) both correspond to different points along this "line of constraints."

The existence of these constraints makes possible a simple formal solution to this class of aperture problems, if measurements made over two or more contours are combined (Adelson & Movshon, 1982). The form of this solution is shown in Figure 19b. Measurements made along the upper right border of the figure ("edge 1") provide one line of constraints; measurements made along the lower right border ("edge 2") provide a second line of constraints; the intersection of these two lines ("object") is the only motion consistent with the two constraints, and *must therefore yield the motion of the object.*

The neural implementation of this model requires that some set of directionally-selective neurons integrates signals from several local measurements of motion. Because of the larger spatial scale of Area MT receptive fields and the fact that Area MT receives directionally selective inputs from Areas V1 and V2, it is natural to suppose that MT might be the site of this integration. It turns out that this is indeed the case. Area MT also contains two distinct kinds of directionally selective neurons. Component direction selective neurons, like neurons in Area V1, provide signals about the local motions of individual contours or orientations. Pattern direction selective neurons, found only in Area MT, carry more fully integrated information about motion that emerges from the combination of signals about motion from several different contours or orientations (Movshon et al., 1986). These neurons provide motion signals that are invariant with the orientation of moving contours and represent a degree of abstraction of motion information not seen at lower levels of the visual pathway.

Our knowledge of the functional characteristics of neurons in the portions of the motion pathway beyond Area MT is relatively sketchy. Even the anatomy is not yet fully understood, and it is likely that areas such as Areas MST (Medial Superior Temporal), 7a and VIP (Ventral Intraparietal) will ultimately prove to have complex functions related to several different aspects of motion processing. For example, some neurons in Areas MST and 7a respond to complex patterns of motion but not to the simple rigid motion of objects across the visual field. Motter and Mountcastle (1981) have shown a pattern of directional responsiveness in parietal neurons that lends itself to an analysis of optic flow produced by locomotion through the environment. More recently, Saito, et al. (1986) reported several complex patterns of response in MST neurons, including preferences for rotations both in the fronto-parallel plane and in depth, as well as for optic flow patterns of the kind suggested by Motter and Mountcastle. Still other data suggest a role for the higher areas of the motion pathway in the control of smooth pursuit eye movements (see, for example, Lisberger, Morris & Tychsen, 1987). Moreover, signals related to motion must also be involved in such basic perceptual tasks as segmentation of complex images (see Nakayama, 1985; DeYoe & van Essen, 1988). Further analyses of this complex and important neural system will surely yield new insights into the brain's processing of visual images.

Motion Perception by a Moving Observer

Up to this point, the chapter has emphasized the dependence of motion perception upon <u>afferent</u> signals. Although such an emphasis is justified, it neglects <u>efferent</u> influences on the perception of object motion. Such influences are clearly important. For example, motion perception is strongly influenced by factors such as concurrent self-motion, eye movements or oculomotor disorders (Brandt & Dieterich, 1988). Although the physiological underpinnings of these effects are not yet understood (Galletti, et al. 1987), they do represent important boundary conditions for the entire field of study.

Under normal everyday conditions, an observer moves freely about within his or her environment. As a result, motion signals arising from the retina subserves two quite different tasks: the observer must control his or her own motion and, at the same time, must also perceive the motion of objects. These two tasks can sometimes be in conflict. For example, while one drives down the road it is difficult to simultaneously perceive movement of roadside tree tops. This fortuitous discovery prompted a series of experiments on object motion perception in the presence of self-motion perception or eye movements. Concern for highway safety lends added importance to the possibility that self-motion and object-motion interact. To take but one example, accurate perception of changes in headway, the distance between cars, is essential to safe driving. If a driver's ability to perceive object motion was impaired by movement of his or her own car, the driver would be much disadvantaged when the situation demanded rapid response to changes in headway.

Mean response times to change in headway, the inter-car separation, have been taken under actual road conditions and compared with measurements made in the laboratory by non-moving observers. Laboratory tests simulated a car's rear end, using an ellipse whose size was varied electronically. Measurements were made with two different reference headways, 20 m (lightly stippled bars) and 40 m (darkly stippled bars). As Figure 20 shows, detection of headway change is much more difficult under actual road conditions (A) than under static laboratory conditions (B and C). The figure also shows that it makes no difference whether the simulation involves an ellipse (B) or a horizontal bar (C) that changes in size. Compared to either case, detection of change is much better than on the road. These results suggest that extrapolations from static, laboratory conditions to predictions of detection on the roadway may underestimate roadway reaction times by as much as several hundred milliseconds (Probst, Krafczyk, Brandt, & Wist 1984; Probst, Krafczyk & Brandt, 1987).

Figure 20 about here

The next experiment deals with the perception of frontoparallel, object motion while the observer's eyes, head or trunk are also in motion. Figure 21 shows that the threshold for detection of object motion increases during concurrent head motion and fixation of the moving target. Data are shown for several different rates of oscillation about the vertical axis, ranging from 0.04 to 0.25 Hz. The oscillations had an amplitude of plus and minus 20°. Notice also that similar results can be obtained without eye or head movement; neck stimulation produced by rotating the trunk relative to the head can also elevate the threshold for object motion (Brandt, 1982; Probst et al., 1986).

Figure 21 about here

Finally, consider the perceptual consequences of certain oculo-motor disorders. The patients whose perceptions will be described had an acquired palsy of the oculomotor, trochlear, or abducens nerve (that is, the third, fourth or sixth cranial nerves). All these patients had difficulties in object motion perception (Brandt & Dieterich, 1986; Dieterich & Brandt, 1987). For example, Figure 22 shows motion perception in a patient's affected and unaffected eyes. The figure also shows the performance of age-matched control observers. The dependent variable is the time required to detect a moving object. Generally, paresis seems to be associated with a suppression of motion perception.

Figure 22 about here

Though suppression of perception of object motion is decidedly abnormal in patients with eye muscle paresis, such suppression confers certain benefits. The perceptual suppression reduces or eliminates the oscillopsia (illusory, perceptual jitter) that would otherwise accompany head movements. This is partly confirmed by the fact that the perceived amplitude of oscillopsia in these patients is always smaller than the net retinal slip. Figure 23 shows that the same holds for patients with an acquired down-beating nystagmus (quick involuntary vertical oscillations of the eye; usually a sign of central nervous system dysfunction) or congenital nystagmus (Dieterich & Brandt, 1987).

Figure 23 about here

Conclusions and Speculations

This chapter has tried to reinforce the notion that the field of motion perception is not completely unified. The diversity of motion becomes particularly clear when one considers the physiological processing of motion information. Different aspects of motion may be processed, or made explicit, at different stages of the magnocellular-stream. As DeYoe & van Essen (1988) emphasize, "motion cues can be used in a diverse range of computational tasks, only some of which are directly related to the perception of object motions per se." Thus, although the magnocellular-stream has often been linked to motion perception, that link --or those links-- may turn out to be quite complex and varied (Van Essen & Maunsell, 1983).

For example, suppose we assert that some particular psychophysical aspect of motion processing emerges at one stage of the magnocellular stream. This assertion assumes what Teller (1984) terms "transparency," the assumption that subsequent stages of the system do nothing to undo this emergent achievement. In other words, those subsequent stages must be transparent. The heterogeneity of neural properties at any one stage of the magnocellular-stream presents quite a different challenge. Many people have recorded from cortical regions that may be involved in motion processing. Nearly every one of those researchers (e.g. Orban, 1986) has commented on the extraordinary cell-to-cell variability in directional selectivity. Because some of the cells at one stage in the magnocellular-stream have properties that parallel human psychophysics in some interesting way, one might be tempted to link psychophysics to the behavior of such cells. However, this fails to account for what the remaining cells in the area are doing, or are not doing. Can we explain how the system filters out those responses? Or do those responses act as a kind of noise?

Researchers interested in motion perception have recently begun to explore these questions. For example, Newsome and Wurtz (1988) used a neurotoxin, ibotenic acid, to create a localized, chemical lesion in Area MT of the monkey. Such lesions produced large, temporary losses in the ability to initiate smooth pursuit eye movements. There is an even more dramatic, hypothetical experiment in which one might ask what everyday vision would be like if one did not have Area MT. This question is interesting, because of a recent clinical report of a patient who purportedly had lost the tissue of Area MT, and neighboring areas (Zihl, von Cramon, & Mai 1983). This patient was extraordinarily impaired on many different tests of movement perception, particularly tests that involved moderate and faster motions (as opposed to very *slow motion). One might assume then, that* loss of Area MT would impair motion perception in a similar manner as the brain lesion impaired performance in this patient. Yet, if a macaque monkey's Area MT is removed bilaterally and in toto, the monkey has no trouble moving around or even responding to objects such as people who move toward the monkey or near it (A. Cowey, unpublished observations).

One possibility is that those signals upon which we base motion perception come through Area MT -when Area MT is available. However, various studies show that if Area MT is destroyed, a relatively short period in which motion perception is severely impaired is followed by a rapid recovery of function. Unfortunately, the patient described by Zihl, von Cramon and Mai (1983) never experienced anything like the recovery that the monkeys do, perhaps because the patient's damage was more extensive.

To return to the first example given at the beginning of the chapter: Exner's demonstration that motion is not merely a derivative from separate analyses of time and space. This demonstration notwithstanding, it is entirely possible that some everyday problems of motion perception could be solved by analyzing where things are and when they are there, without actually making the motion signal explicit. In other words, target motion does not in and of itself guarantee that the target is processed by a specialized motion-processing system. For example, when a target moves extremely slowly, does perception of that target necessarily depend upon the special "machinery" normally involved with motion perception? A major challenge for future research is learning to define the conditions under which Exner was right.

Figure Captions

Figure 1. Basic scheme of Reichardt's motion detector. Panel A: Signals from two photoreceptors (shaded rectangles) are sent to unit M where the signals arriving at various instants are multiplied. The signal from the left receptor reaches the multiplier after some delay, Δ, relative to the signal from the right receptor. As a result, the multiplier unit responds well to a pattern moving in the direction of the arrow. A stimulus would be particularly effective if it first stimulated the left receptor and then, with delay Δ, stimulated the right receptor. Under these conditions, the product of the two signals would be large, as would be the output of unit M. Panel B: A second group of photoreceptors and associated multiplier unit. The position of the delay, Δ, makes this group respond poorly or not at all to rightward motion. Its preferred direction of motion, shown by the arrow, is leftward. Panel C: A more complete Reichardt unit, with two pairs of receptors, multiplier units and delays. The left multiplier unit would respond well to a pattern that travels across the retina from left to right (with appropriate velocity); the right multiplier would do the same for a pattern travelling from right to left (again, with appropriate velocity). A final, subtraction unit, not shown, would convert the difference between the two M units' outputs into a directional response.

Figure 2. Each panel shows three frames of a simple random dot cinematogram. Some dots have been shifted from the first frame (top) to the second frame (middle) and then to the third frame (bottom). If the frames were presented as a random dot cinematogram, the shifted dots would immediately manifest themselves in apparent movement. In the left panel, some effort is required to identify the shifted dots, in still frames. In the right panel, the three frames of the cinematogram are shown with the shifted dots highlighted for ease of identification.

Figure 3. Construction of a random-dot cinematogram. Two patterns of random dots are presented in rapid succession: a typical row from the first and second pattern is shown. In a random-dot cinematogram of the kind shown, only the dots within a central region undergo a coherent displacement, and the subject is asked to report the shape (vertical or horizontal) of this region. An alternative method is to displace all the dots in the pattern, requiring the subject to report the direction of motion. (From Braddick, 1974).

Figure 4. Variation of d_{max} with eccentricity, e, in degrees. The display consisted of dots displaced upwards or downwards within two vertical strips. The width and length of the strips are scaled as e changes so that the width always equal to e/3 and length always equals 2e. Moreover, the outer edges of the vertical strips are maintained at distance e on either side of the fixation point.(After Baker & Braddick, 1985)

Figure 5. Directional judgments for displacements in narrowband (0.5 octave) spatially filtered random dot cinematograms. The three curves are data from patterns whose center spatial frequencies are in the ratio 1:2:4. Displacements are plotted on the x-axis as multiples of the period of the center frequency for each pattern. d_{max} is taken as the displacement for which the error rate (y-axis) reaches 20% on the first rising part of the curve. The error rate exceeds 50% when the displacement size is between 25 and 50% the period of the pattern (From Cleary, 1988)

Figure 6. d_{max} as a function of the interval (ISI.) between dot pattern exposures, for three different eccentricities with each of two subjects. For ISI values higher than the rightmost point of each curve, direction of motion could not be reported for any size displacement, so d_{max} can

be taken as zero. Each pattern was exposed for 60 msec. (From Baker & Braddick, 1985.)

Figure 7. Maximum displacement for directional response in macaque cortical neurons. Triangles and dashed regression line: cells in Area V1. Circles and solid regression line: cells in Area MT. The solid squares show human psychophysical data for comparison, taken from the results illustrated in Figure 4. (From Mikami, Newsome & Wurtz, 1986).

Figure 8. Increase of d_{max} with increasing number of displacements in a sequence: d_{max} as plotted refers to the size of each individual displacement in the sequence. The different symbols refer to different rates of presentation: the value given in the legend is the stimulus onset asynchrony, i.e., the time between onset of successive pattern exposures. (From Snowden & Braddick, 1987.)

Figure 9. Ratio of receptive field (RF) width to maximum displacement for directional response, for neurons in macaque Area MT. The plotted regression line indicates the shallow increase of this ratio with receptive field eccentricity. (From Mikami, Newsome, & Wurtz, 1986.)

Figure 10. An arrangement for demonstrating rotary inertia. The top row illustrates the two-frame sequence in a control condition. The display alternates between two, spatially-overlapping crosses that are rotated 45 degrees relative to one another. The resulting apparent motion is ambiguous, rotation is perceived either clockwise or counterclockwise. The bottom row illustrates a three-frame sequence designed to show inertia. Note that a new, third cross has been added before the original sequence. The relative orientations of the first and second targets produce strong clockwise motion. This strong motion persists when the third element is

presented and the sequence is repeated.

Figure 11. A: Display whose first frame comprises a triangle and a square, and whose second frame comprises just a single square located adjacent to the position previously occupied by the triangle. B: Diagram of the percept produced by the display. The square appears to move obliquely upward and to the right; the triangle appears to move horizontally, ultimately being occluded by the square.

Figure 12. Lower panel: a quartet of discs that produces a randomly fluctuating percept. When the pair of lighter discs (labelled "1") is presented in alternation with the pair of darker discs (labelled "2") the percept varies randomly. The discs will seem to move either up and down or left and right (as indicated by the arrows). Upper panel: when many quartets are presented in an array, the random fluctuations of individual quartets seem to be synchronized: at one time all seem to move up and down, at another time all seem to move left and right.

Figure 13. Responses of cortical cells (cat Area 17) to opposite direction of bar motion as a function of texture motion. The texture was either stationary (0), or moved in the left or right direction, at the same speed as the bar (sa), four times slower (sl) or four times faster (fa). The dotted horizontal lines indicate the significance level; an asterisk indicates a response in the preferred direction that is significantly different from the responses to that direction in the control condition (texture stationary). The neuron in A remains direction selective for all texture motion conditions; the neuron in B loses direction selectivity when the texture moves in phase with the bar. From Orban, Gulyas & Vogels (1987).

Figure 14. Responses of a velocity tuned cell (cat Area 18) to light and dark bars moving in a preferred direction (to the right) or in a non-preferred direction (to the left). The cell is tuned to the same speed (9.5ß/sec) for light and dark bars. For both bars, the cell is not direction selective at 1ß/sec, at slow and high speeds, but becomes completely direction selective at medium speed. (From Gulyas, Lagae & Orban, unpublished).

Figure 15. Percentage of cells plotted as a function of optimal velocity for cells in Areas 17, 18 and 19 of the cat (A) and of Area MT cells with strong velocity tuning in the macaque (B). Distributions are plotted for 3 ranges of eccentricities (as indicated on the left of the histograms). The third range of eccentricity shown at the bottom extended up to 35ß in the cat and up to 25ß in the monkey. The data in B are plotted from Maunsell & Van Essen (1983b). From Orban (1985).

Figure 16. Three dimensional plots of response rate in the preferred direction as a function of apparent-velocity and strobe rate of a cat Area 17 velocity tuned cell. The dashed parts of the velocity-response curves indicate responses below the significance level. Horizontal thin lines indicate mean spontaneous activity. From Cremieux, Orban amd Duysens (1984).

Figure 17. Just noticeable differences in perceived velocity expressed as Weber fractions ($\Delta\omega/\omega$) and plotted against stimulus velocity (ω). Data are averages from two human subjects. Light bar continuously illuminated (filled circles), light bar of low luminance (reduced 10 fold) stroboscopically illuminated at 100 Hz (crosses), and light bar of high luminance stroboscopically illuminated at 10 Hz (circles). The loss in velocity discrimination at 10 Hz is not due to a reduction in total energy. (DeBruyn & Orban, unpublished).

Figure 18. Post-stimulus time histograms (PSTHs) representing the average response (n=20) to a light bar moving horizontally over the texture (A), to the same bar moving in an opposite direction over the texture (B), and to the texture moving on its own (C). This cell was recorded from layer 6 in Area 17 of the cat and its receptive field was centered 3.4ß from fixation. Each row of PSTHs corresponds to a background condition indicated by a number between 1 and 7. Conditions 1 to 3 correspond to texture motion to the left, condition 4 corresponds to a stationary texture, and conditions 5 to 7 to the texture moving to the right. In conditions 2 and 6, the texture moves at the same speed as the bar (2.2ß/s); in conditions 3 and 5, slower than the bar (0.5ß/s), and in conditions 1 and 7, faster than the bar (8.8ß/s). From Orban, Gulyas & Vogels (1987).

Figure 19. The aperture problem. **A.** Two diamonds, one moving downward and one moving to the right, showing that locally measured motions (circles) do not unambiguously reflect the overall motions of objects. **B.** One formal solution to the aperture problem based on using the intersection of the constraints set up by local measurements to resolve their ambiguity.

Figure 20. Object-motion perception under actual road and simulated conditions. Mean response times were determined for the perception of changes in headway at distances of 20 m (lightly stippled bars) and 40 m (darkly stippled bars) under actual conditions (A) and simulated conditions without concurrent self-motion (B and C). In measurements for actual road conditions, the subject was in a moving car. An approximation of the perceptually effective area of the rear of the leading car was simulated by an electronically-generated ellipse

of equivalent retinal size. Headway changes were simulated by adjusting the retinal ellipse area. The times to detect changes in headway were significantly higher for the actual road condition. Under static conditions in the laboratory there was no difference between the detection of a gradual change in area of the ellipse (B) and a horizontal bar with the same but one-dimensional movement (C). (After Probst, Krafczyk, Brandt, & Wist 1984).

Figure 21. Object-motion perception with head or trunk oscillations. Mean response times (in msec) plotted as a function of oscillation. Target speed was 5 deg/sec. There were three modes of simultaneous body-motion. The target was fixated during horizontal head oscillation with vestibular-ocular reflex (VOR) (A), or with fixation suppressd (B), or with the head fixed by the helmet and pure cervical stimulation provided by trunk oscillations (C). Abscissa shows different frequencies of oscillation. Response time to detect object-motion increases with increasing frequency of either head or trunk oscillations.

Figure 22. Response times (means and standard deviations) required to detect horizontal object motion as a function of subject age. The stimulus moved at a constant angular velocity of 24 min/sec. The shaded areas represent results from a control group, 60 neurological patients without ocular motor disturbances (n=10 for each decade from 10 to 70 years). Response times are shortest at about 20 years of age with increasing mean values and standard deviations in the elderly. For comparison, the data from 27 patients with acquired extraocular eye muscle pareses are shown. These patients exhibit longer response times for monocular vision with the affected eye (filled circles) as well as the normal eye (open circles). Impairment of motion perception is more pronounced for the affected eye.

Figure 23. Object-motion perception as a function of the eccentricity of horizontal gaze in patients with congenital nystagmus and acquired downbeat nystagmus. Thresholds for detection of object-motion (24 min/sec; means and standard deviations) as a function of the eccentricity of horizontal gaze in patients suffering from congenital nystagmus and acquired downbeat nystagmus as compared to normals. Thresholds are indicated, on the left ordinate, as ΔT (exposure time in seconds) or, on the right ordinate, as DS (displacement of stimulus in min of arc). Normals first show only a slight increase in thresholds with eccentric gaze and then show a more pronounced increase on lateral gaze of 40 deg. However, whether the ocular oscillation is congenital or acquired, the patients' thresholds are significantly raised. There is a large increase in threshold for directions of gaze beyond 20 deg. As a result the amplitude of the nystagmus is increased.

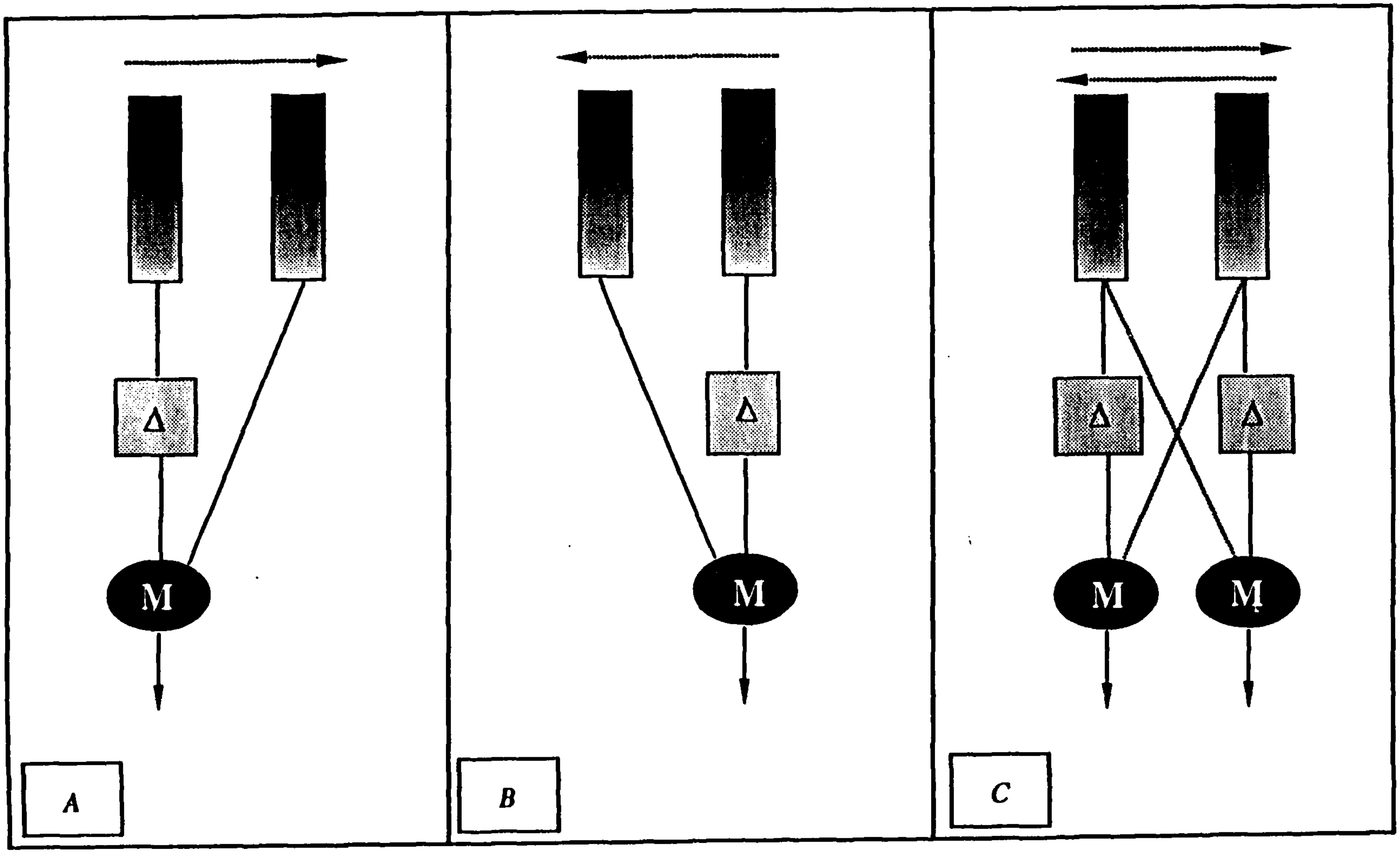
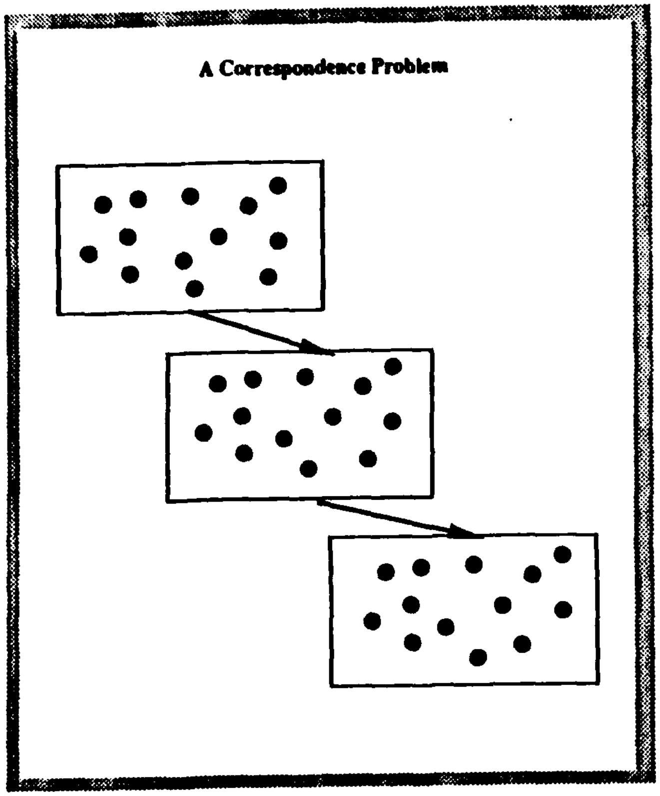
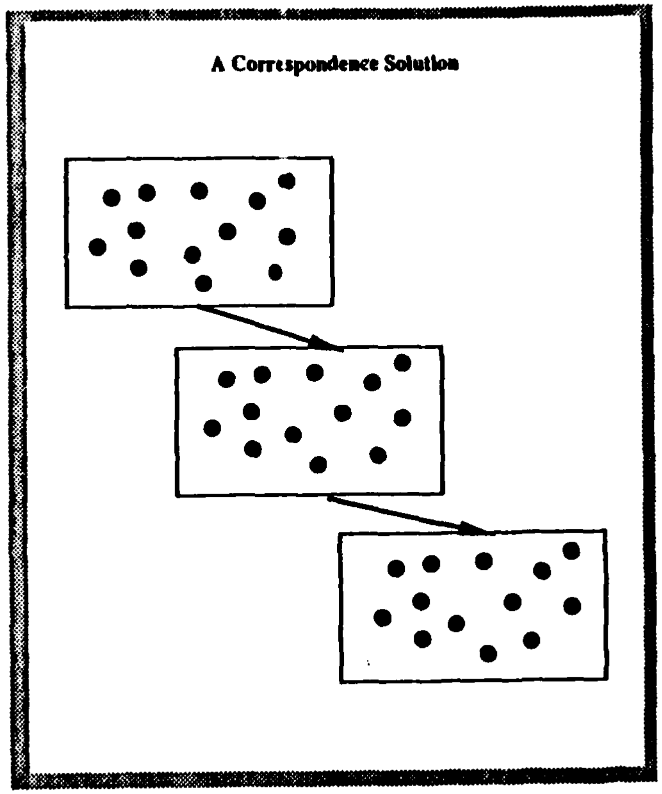


Figure 1


A Correspondence Problem


A Correspondence Solution

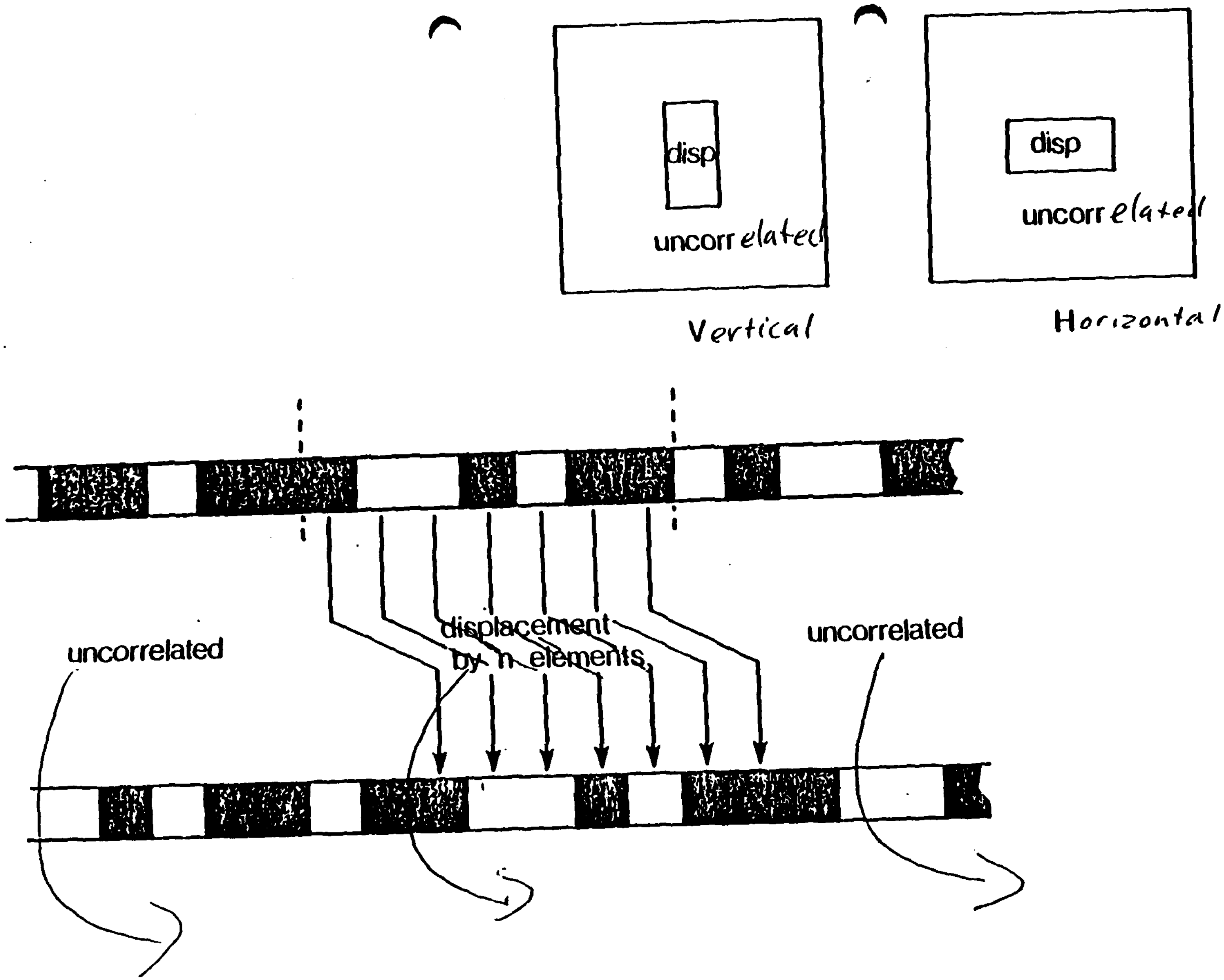

disp
uncorrelated
Vertical
disp
uncorrelated
Horizontal
displacement
by n elements
uncorrelated
uncorrelated

Figure 3

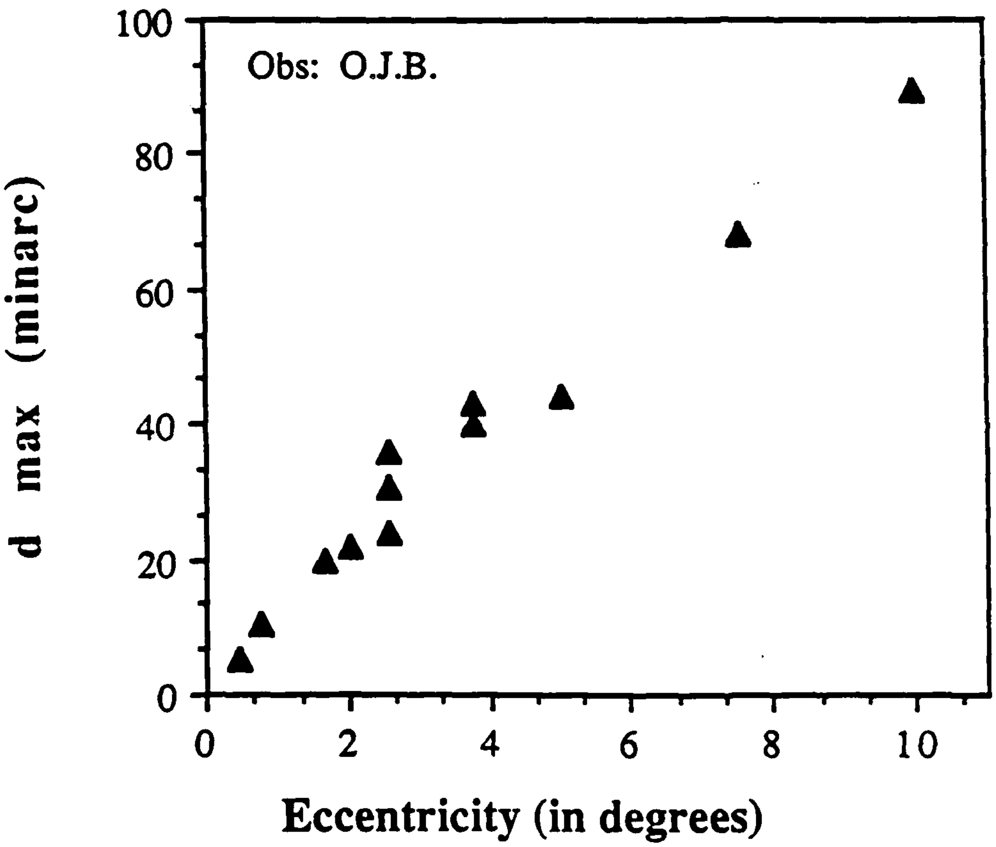
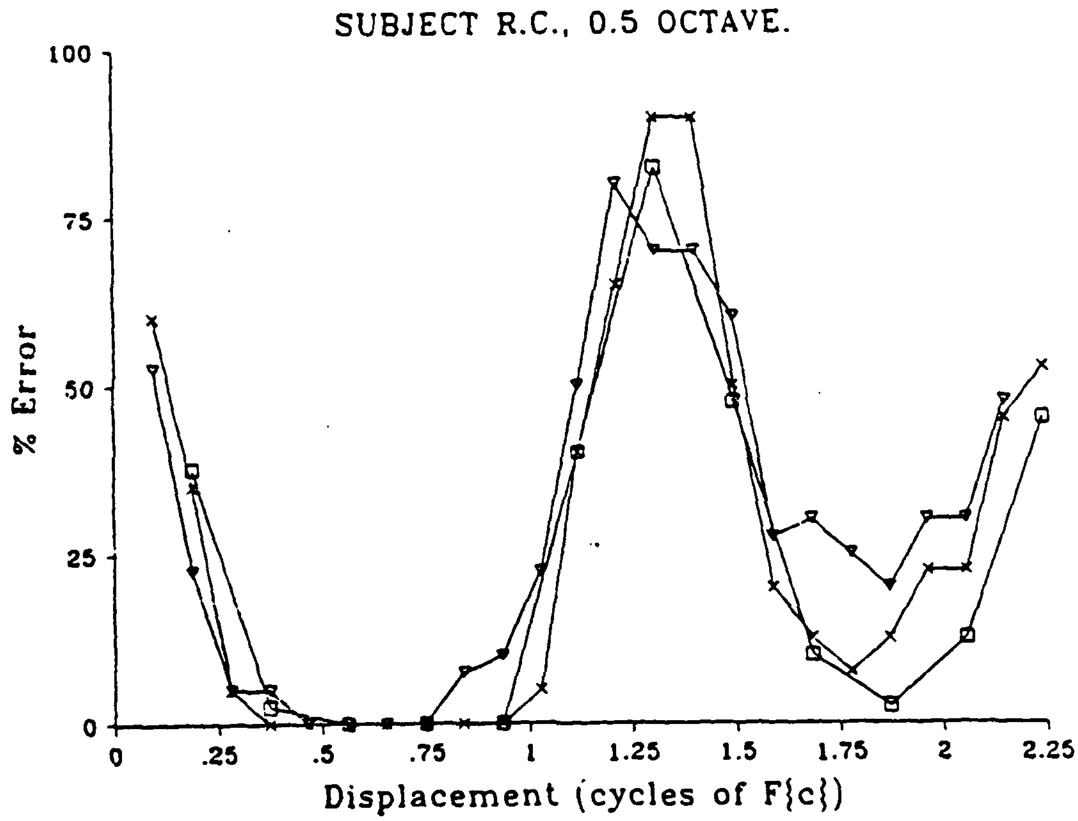

Obs: O.J.B.
d max (minarc)
100
80
60
40
20
0
0
2
4
6
8
10
Eccentricity (in degrees)

Figure 4



SUBJECT R.C., 0.5 OCTAVE.
100
75
50
25
0
% Error
0
.25
.5
.75
1
1.25
1.5
1.75
2
2.25
Displacement (cycles of F{c})

Figure 5

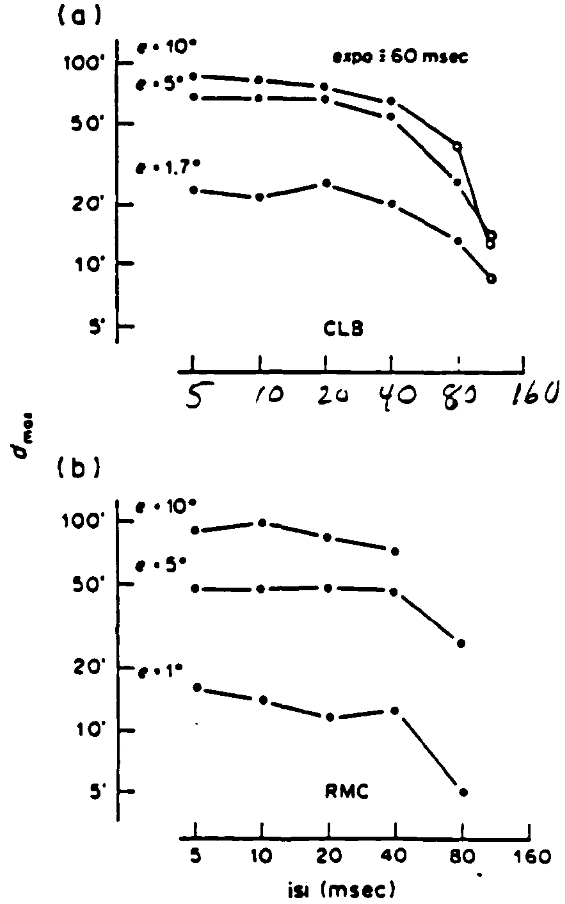

(a)
e = 10°
e = 5°
e = 1.7°
expo = 60 msec
100'
50'
20'
10'
5'
CLB
5
10
20
40
80
160
d_{max}
(b)
e = 10°
e = 5°
e = 1°
RMC
isi (msec)

Figure 6

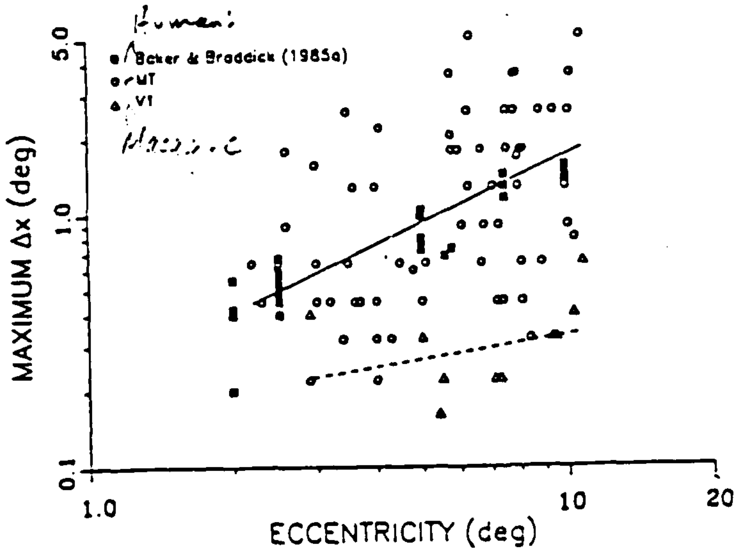
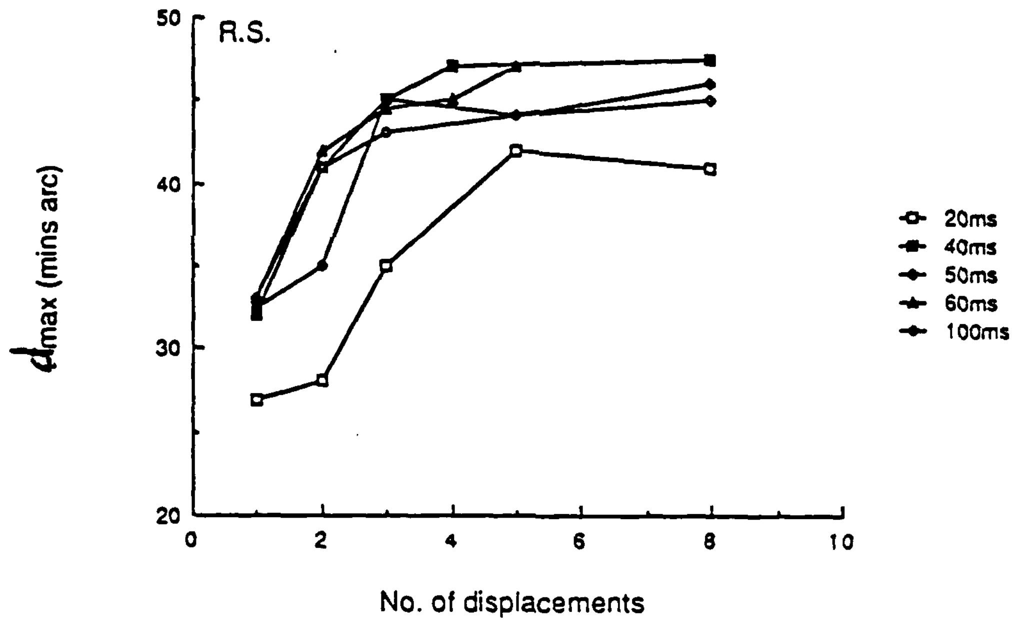

Baker & Braddick (1985a)
MT
V1
5.0
1.0
0.1
MAXIMUM Δx (deg)
1.0
10
20
ECCENTRICITY (deg)

Figure 7


R.S.
dmax (mins arc)
No. of displacements
20ms
40ms
50ms
60ms
100ms

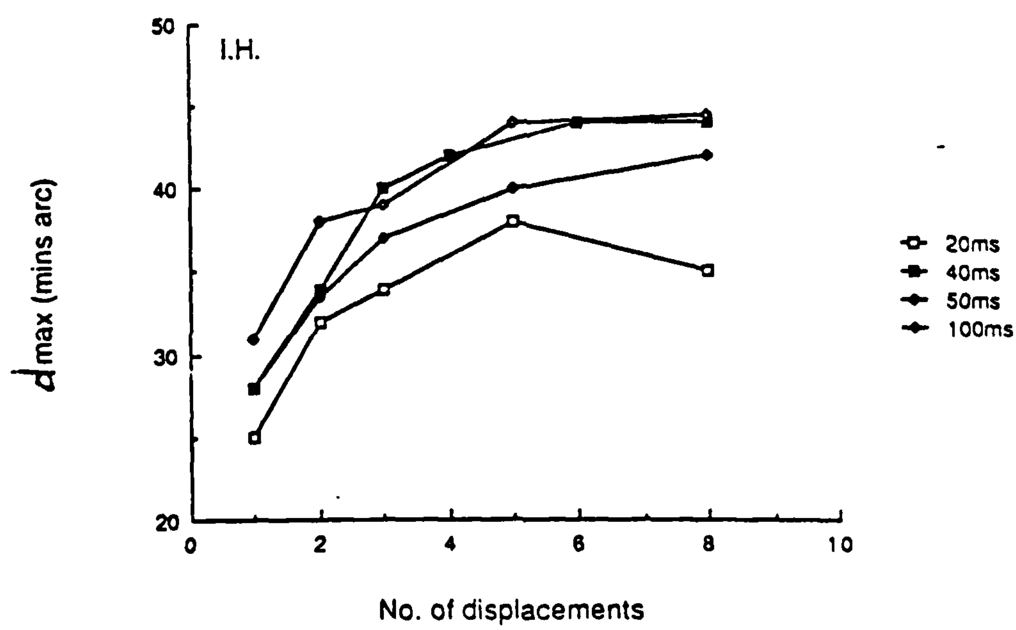

I.H.
dmax (mins arc)
No. of displacements
20ms
40ms
50ms
100ms

Figure 8

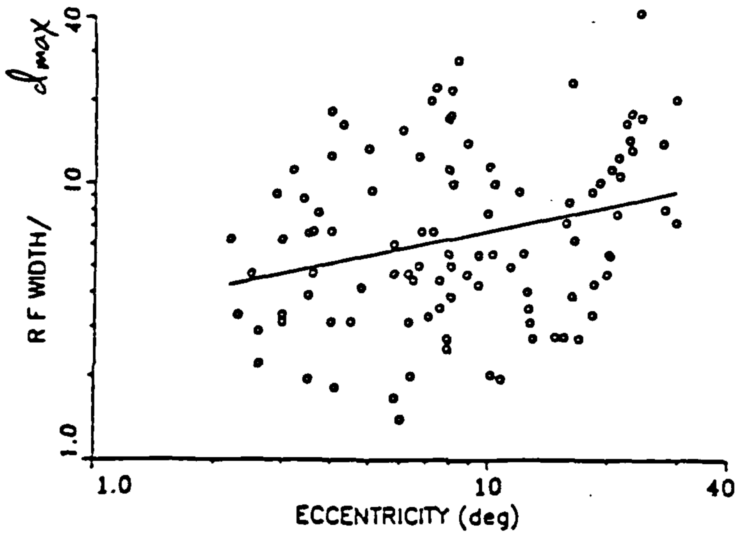


Figure 9

Please write numbers
on vertical axis.

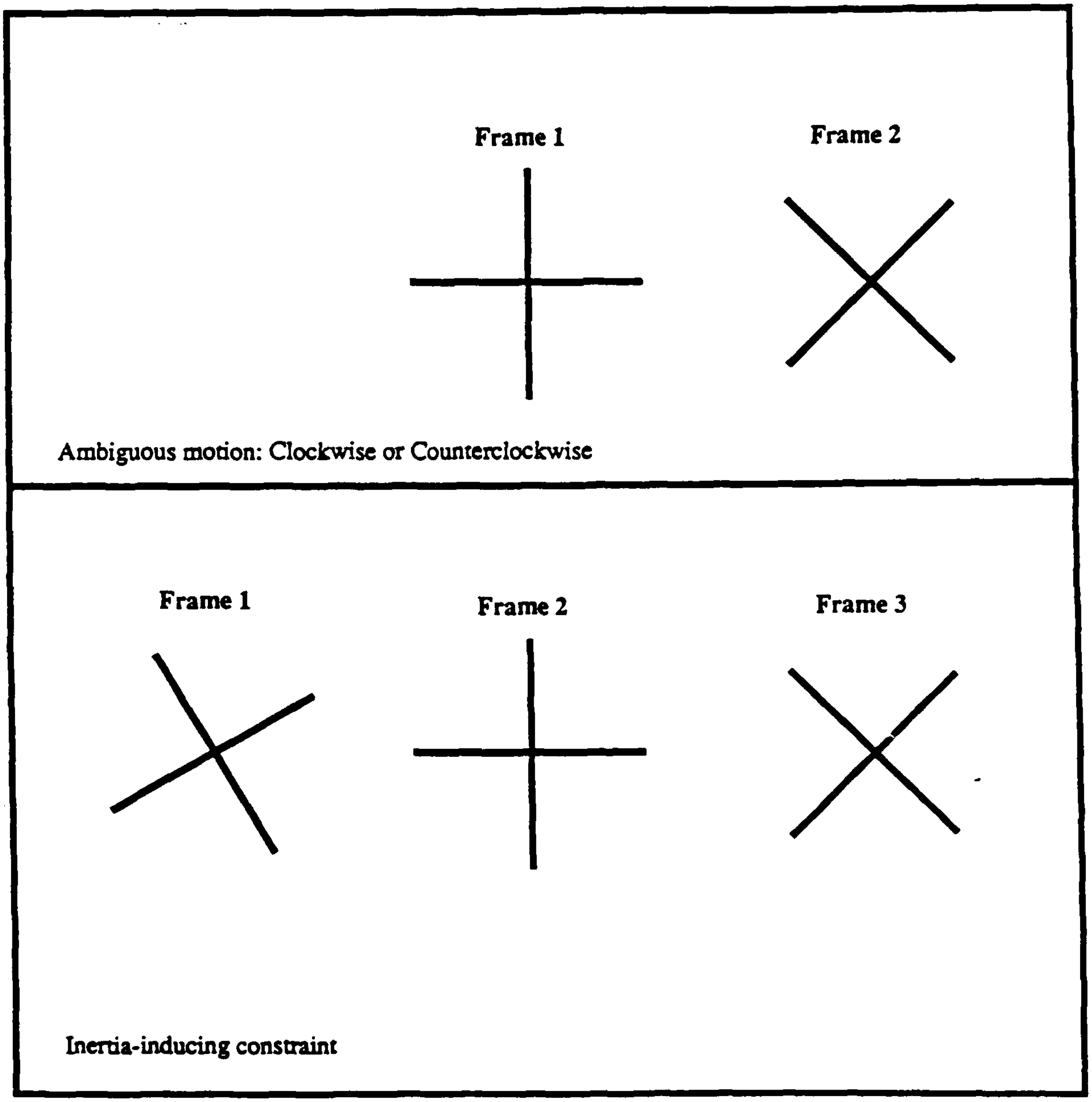

Frame 1
Frame 2
Ambiguous motion: Clockwise or Counterclockwise
Frame 1
Frame 2
Frame 3
Inertia-inducing constraint

Figure 10

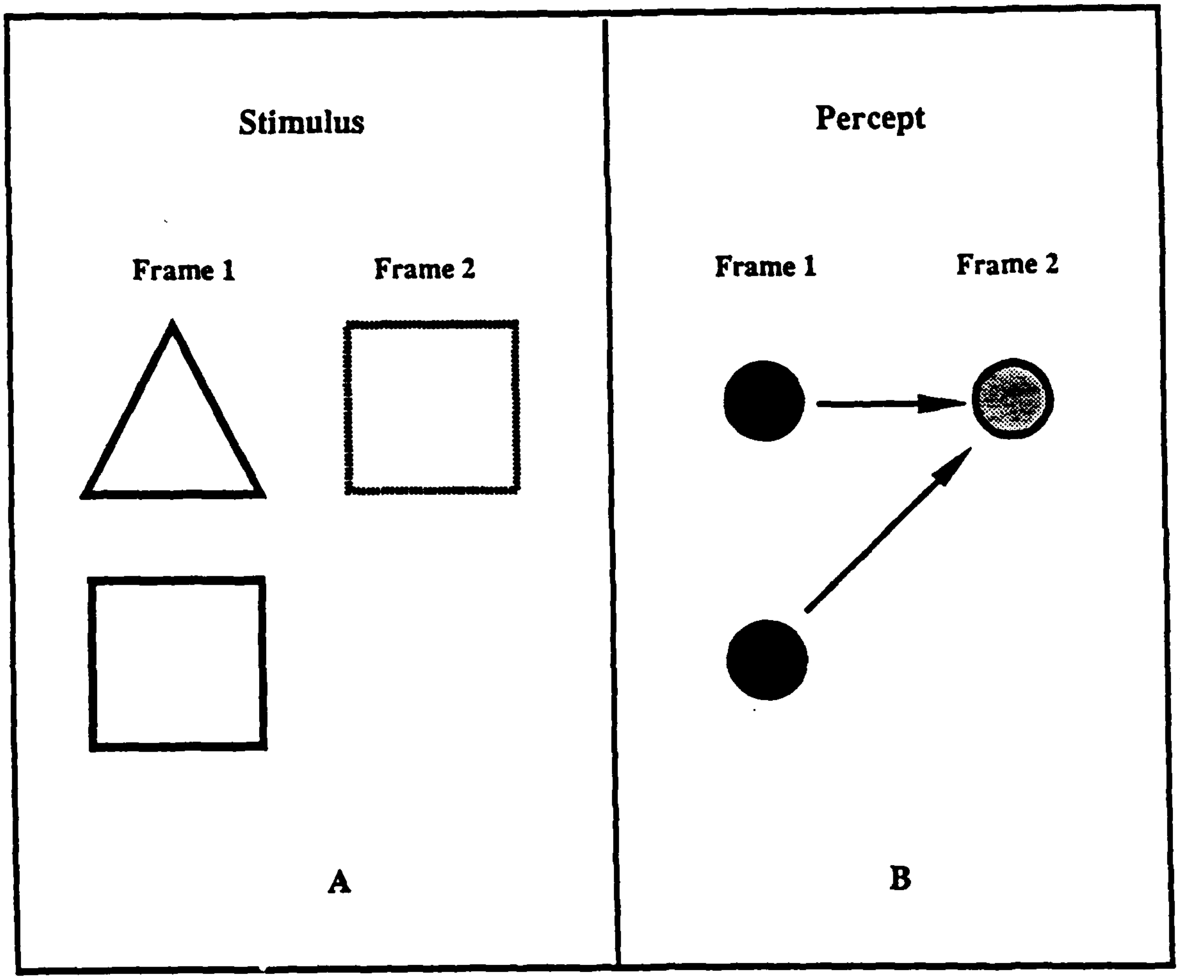
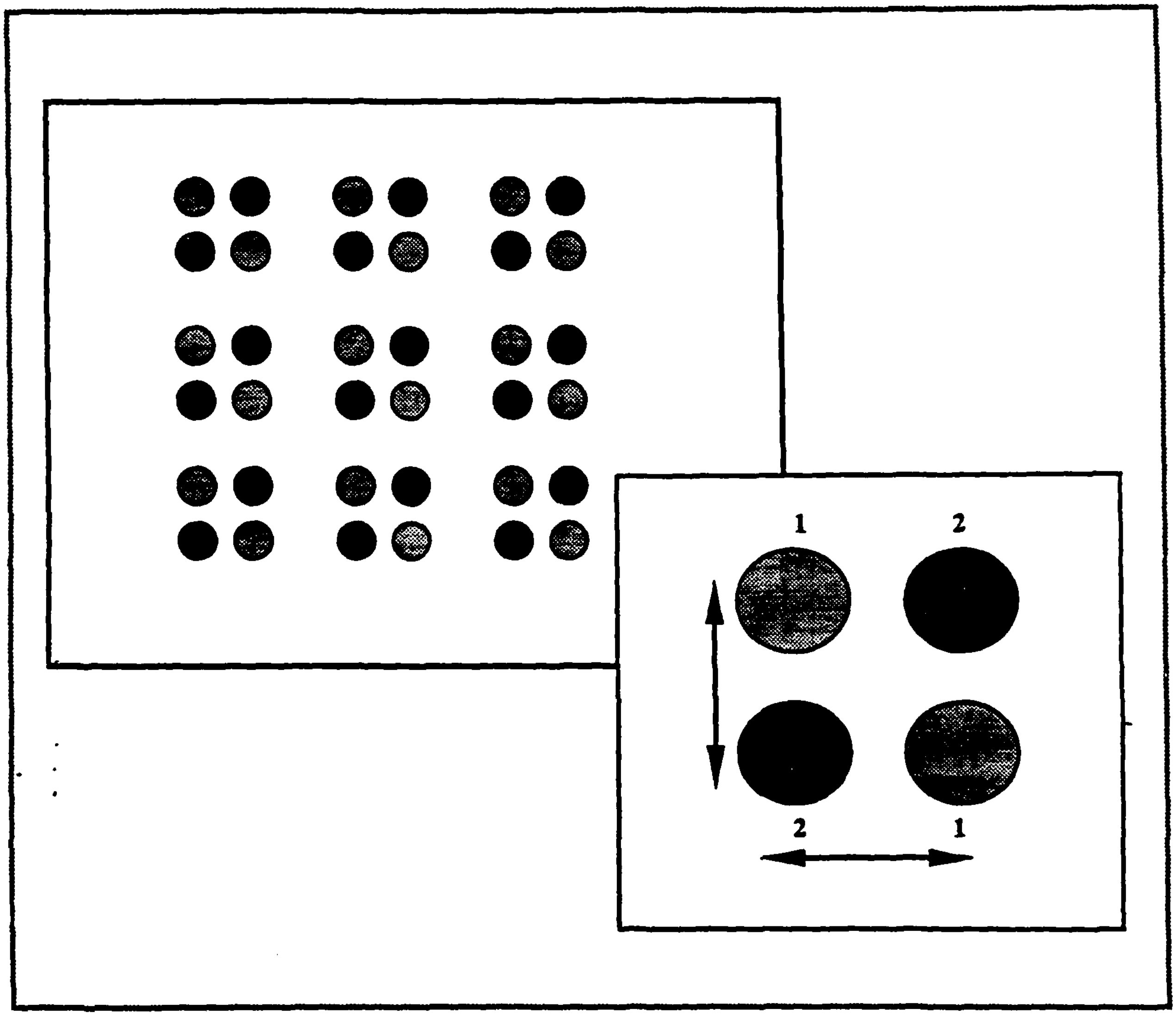


Figure 11



261

Figure 2

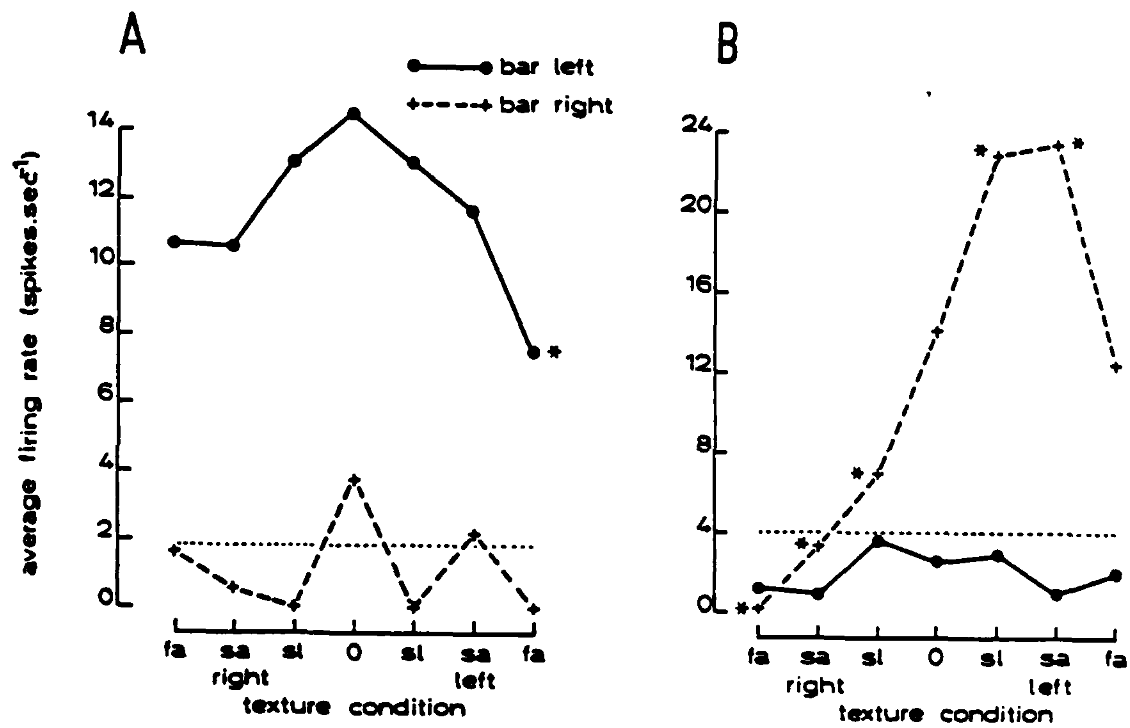
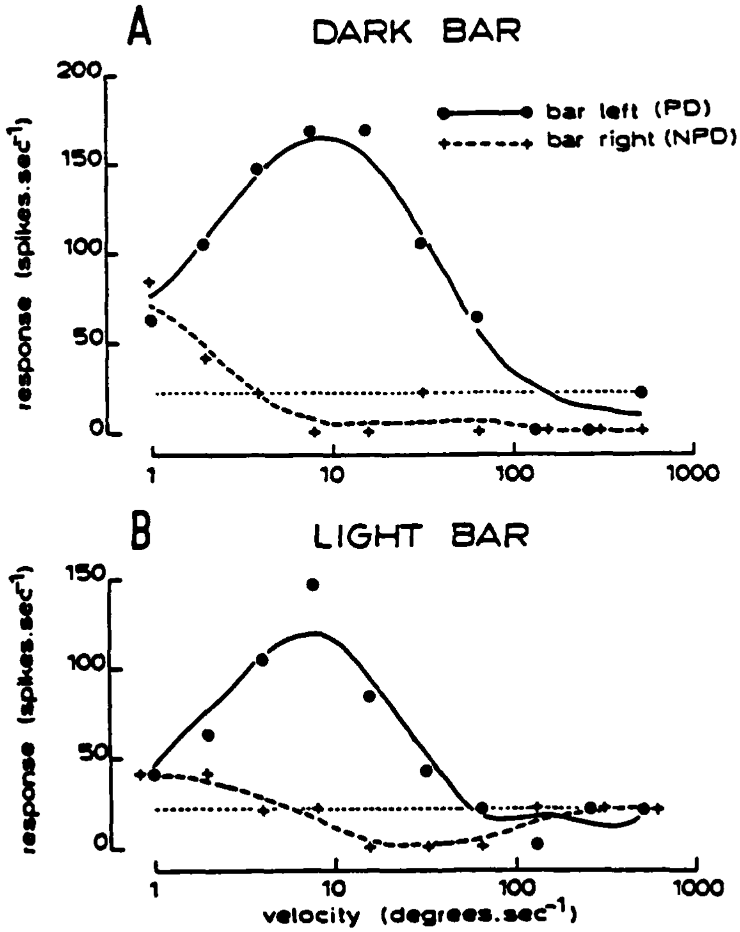

A
B
bar left
bar right
average firing rate (spikes.sec⁻¹)
fa
sa
sl
0
sl
sa
fa
right
left
texture condition

Figure 13



A
DARK BAR
200
150
100
50
0
response (spikes.sec^{-1})
bar left (PD)
bar right (NPD)
1
10
100
1000
B
LIGHT BAR
150
100
50
0
response (spikes.sec^{-1})
1
10
100
1000
velocity (degrees.sec^{-1})

Figure 14

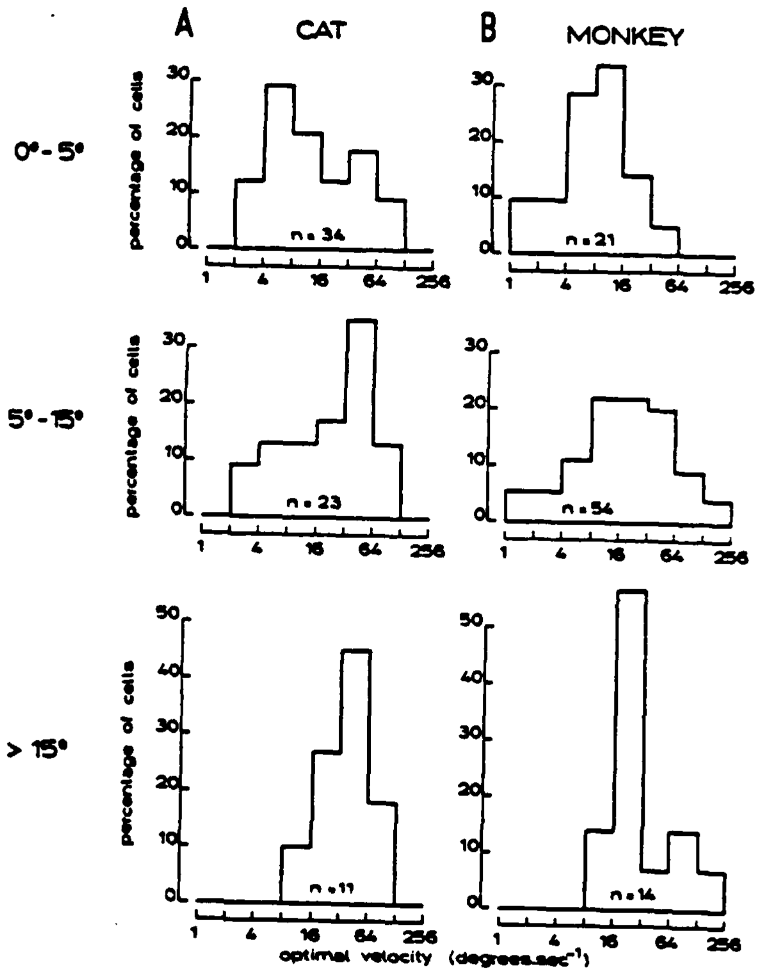
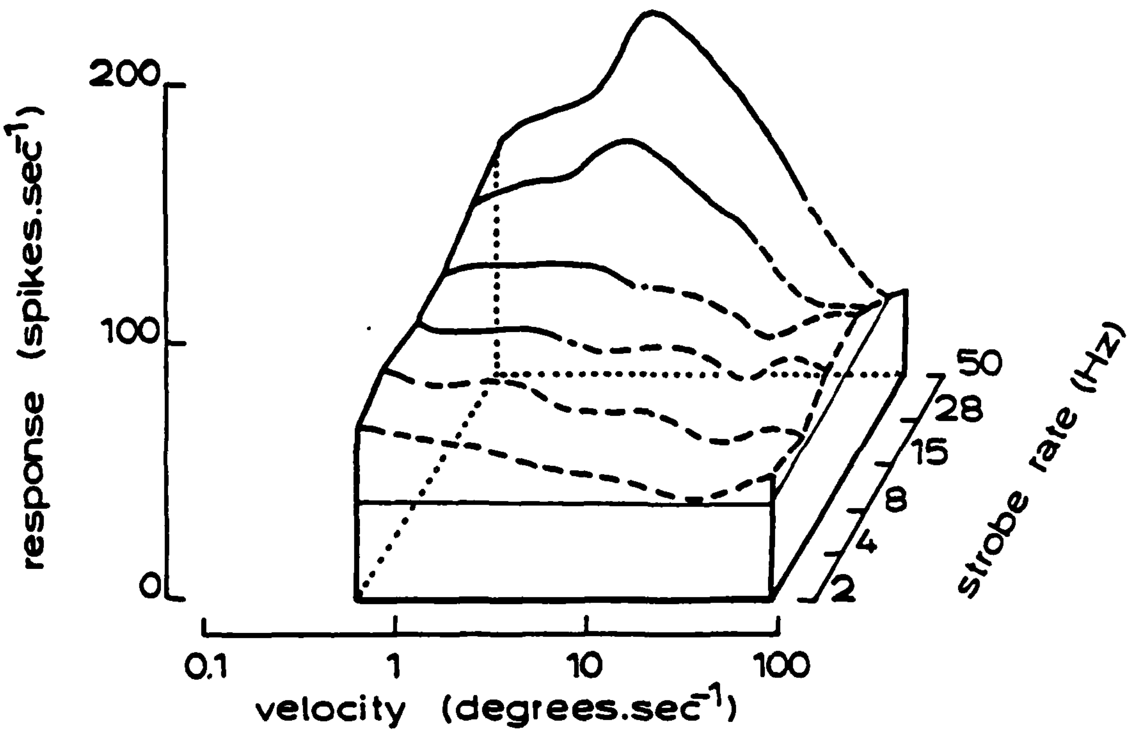

A
CAT
B
MONKEY
0°-5°
5°-15°
> 15°
percentage of cells
n = 34
n = 21
n = 23
n = 54
n = 11
n = 14
optimal velocity (degrees.sec^{-1})

Figure 15



200
100
0
response (spikes.sec^{-1})
0.1
1
10
100
velocity (degrees.sec^{-1})
50
28
15
8
4
2
strobe rate (Hz)

Figure 16

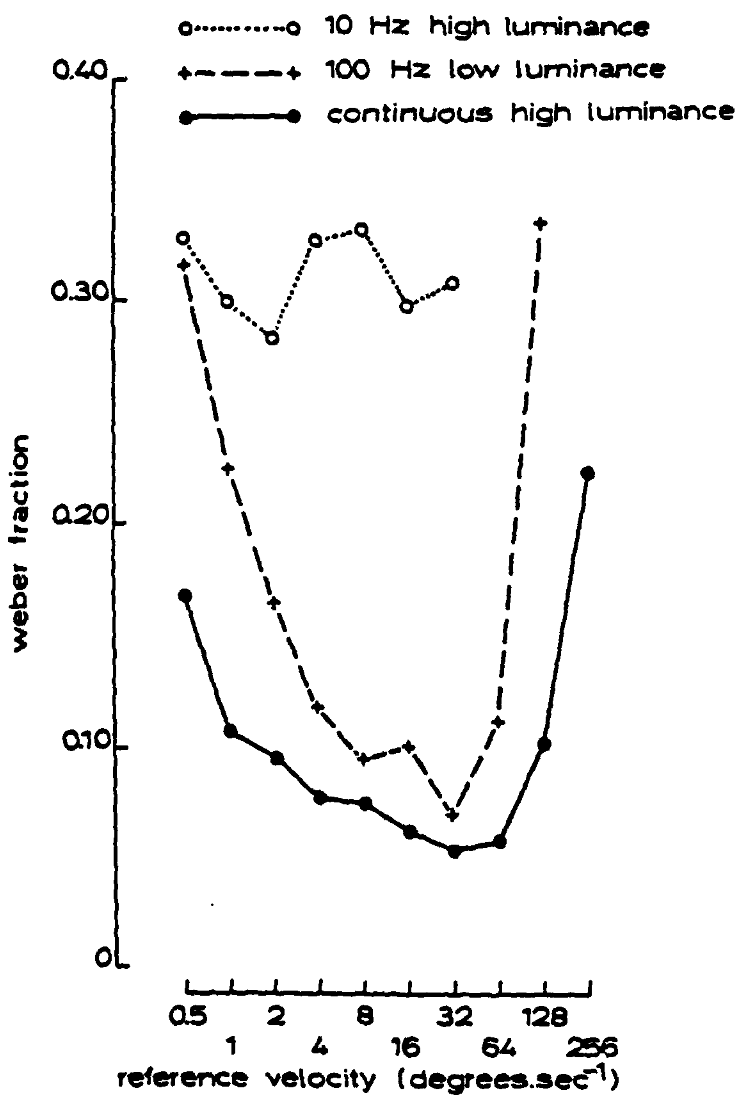

10 Hz high luminance
100 Hz low luminance
continuous high luminance
0.40
0.30
0.20
0.10
0
weber fraction
0.5
1
2
4
8
16
32
64
128
256
reference velocity (degrees.sec^{-1})

Figure 17

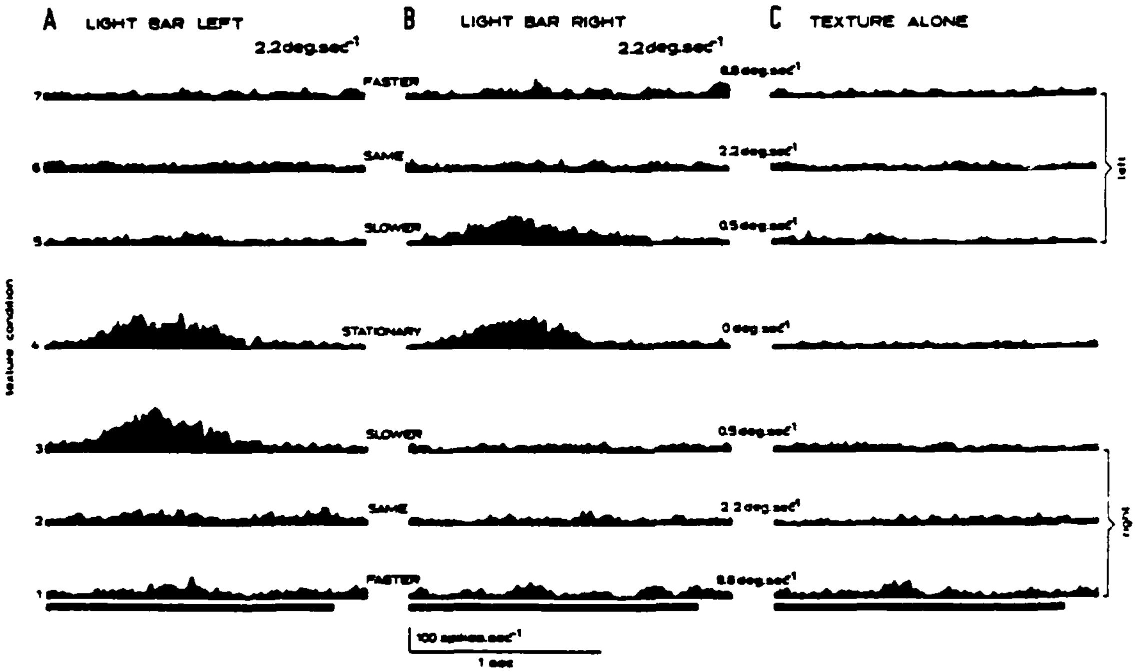
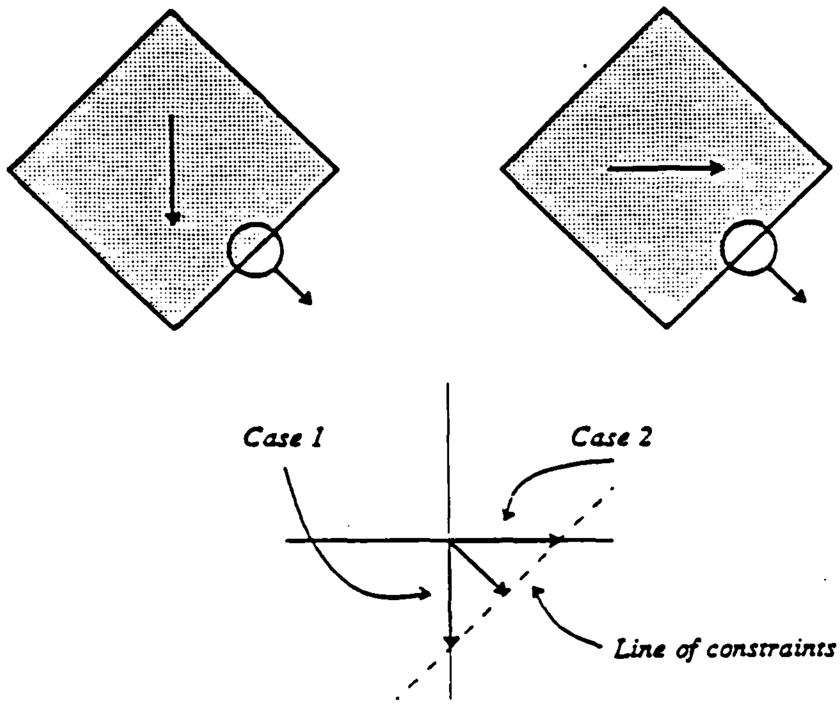
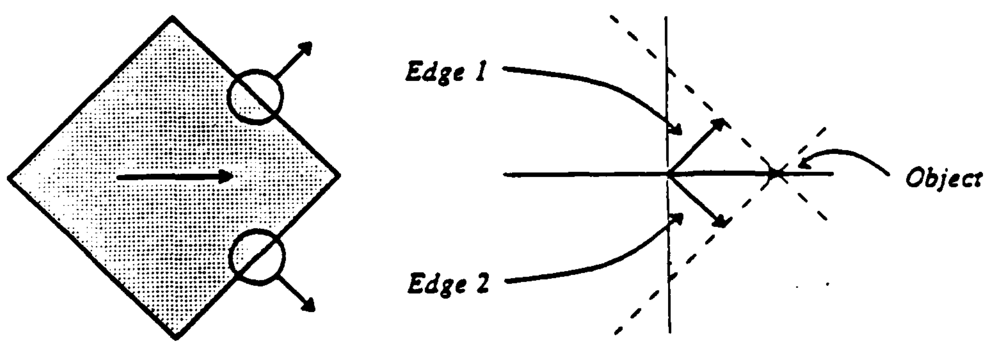

A LIGHT BAR LEFT
2.2deg.sec-1
B LIGHT BAR RIGHT
2.2deg.sec-1
C TEXTURE ALONE
FASTER
SAME
SLOWER
STATIONARY
SLOWER
SAME
FASTER
8.8 deg.sec-1
2.2 deg.sec-1
0.5 deg.sec-1
0 deg.sec-1
0.5 deg.sec-1
2.2 deg.sec-1
8.8 deg.sec-1
texture condition
left
right
100 spikes.sec-1
1 sec

Figure 18

A



B



268

Figure 19

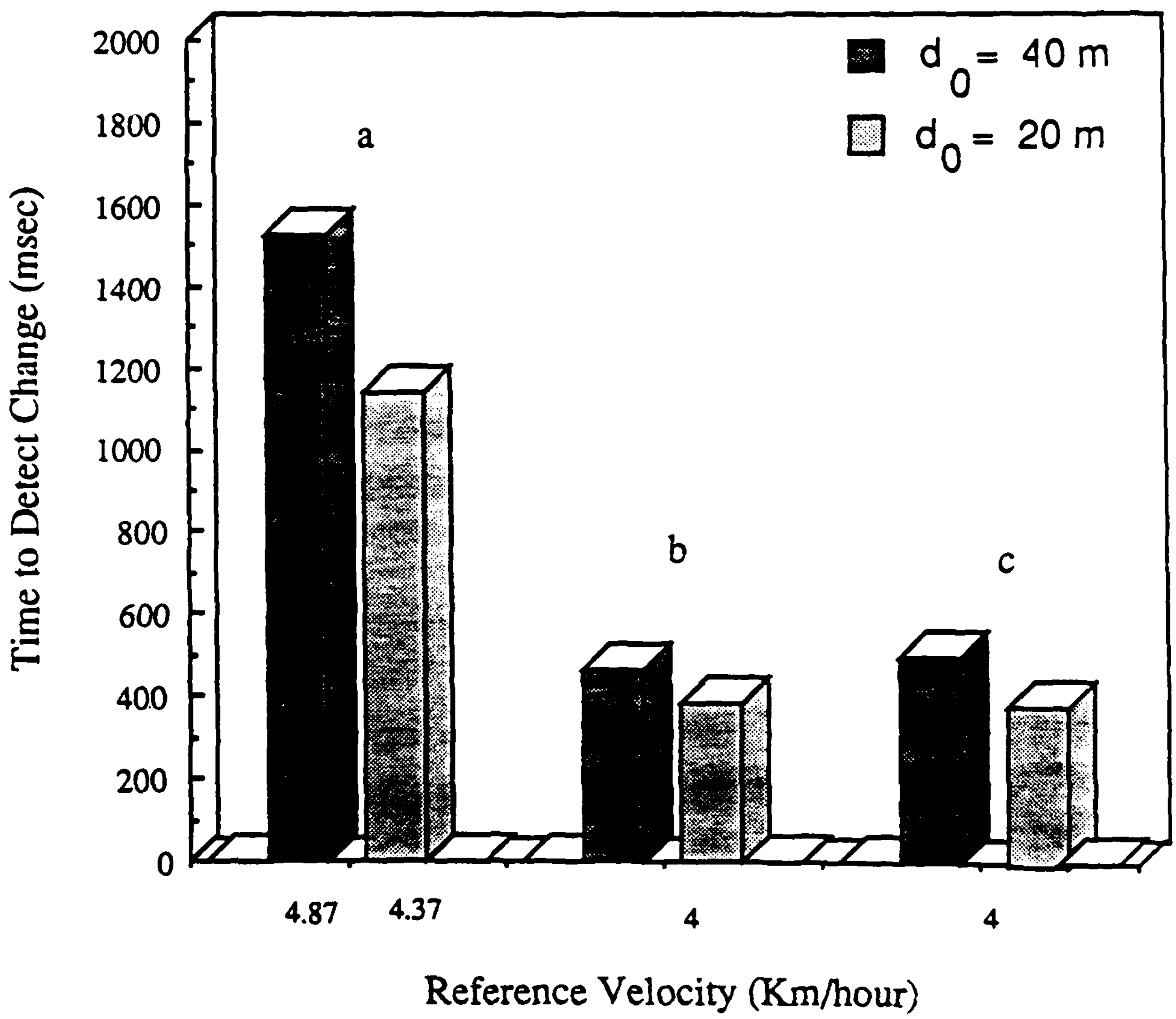

2000
1800
1600
1400
1200
1000
800
600
400
200
0
Time to Detect Change (msec)
a
b
c
d0 = 40 m
d0 = 20 m
4.87
4.37
4
4
Reference Velocity (Km/hour)

Figure 20

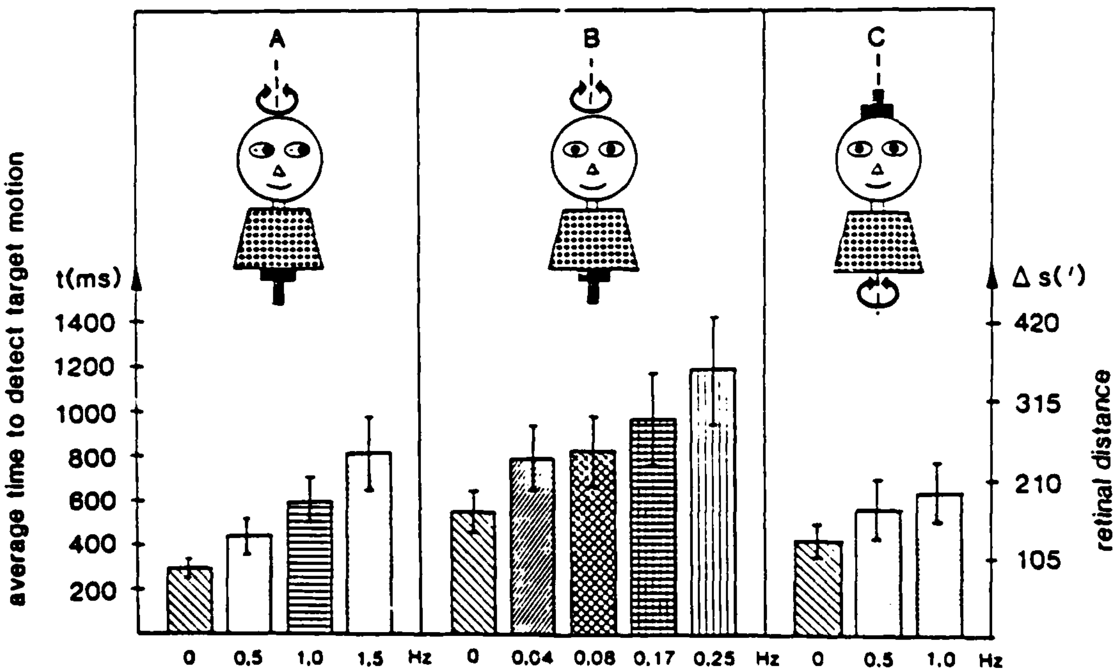

A
B
C
t(ms)
1400
1200
1000
800
600
400
200
average time to detect target motion
Δ s(′)
420
315
210
105
retinal distance
0
0.5
1.0
1.5
Hz
0
0.04
0.08
0.17
0.25
Hz
0
0.5
1.0
Hz

Figure 21

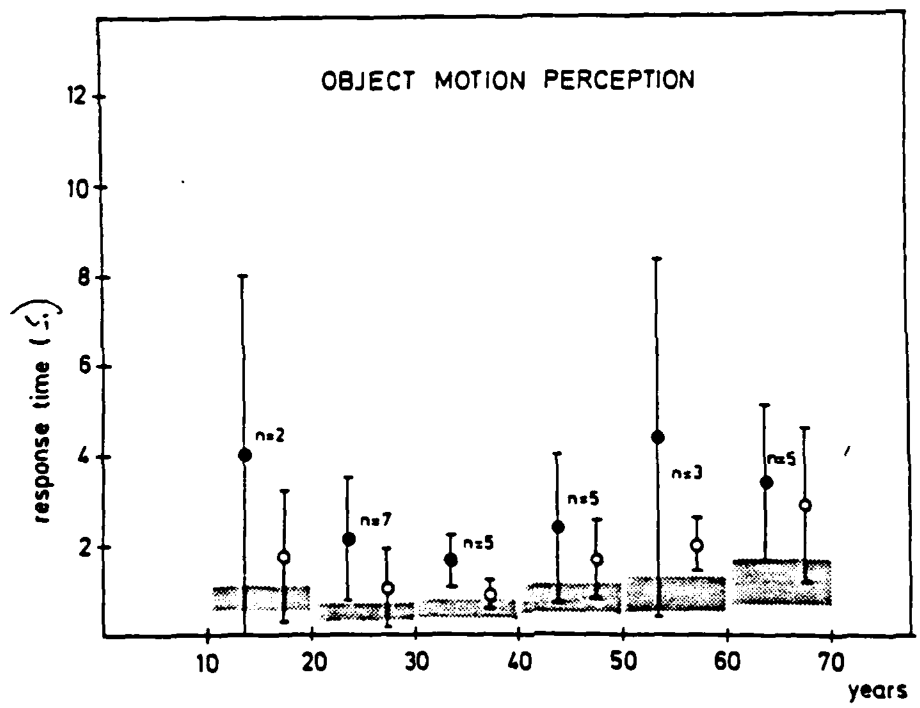

OBJECT MOTION PERCEPTION
response time (s)
12
10
8
6
4
2
10
20
30
40
50
60
70
years
n=2
n=7
n=5
n=5
n=3
n=5

Figure 22

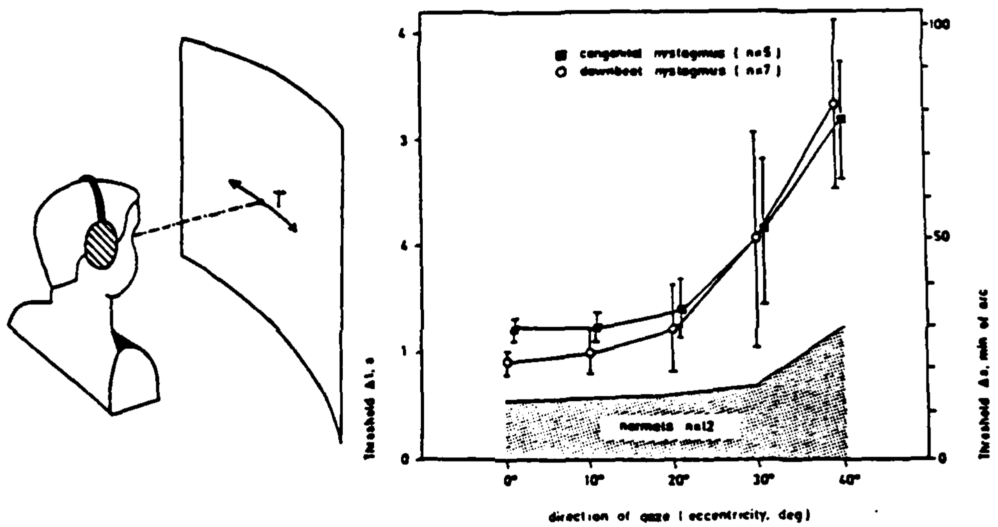

T
congenital nystagmus (n=5)
downbeat nystagmus (n=7)
normals n=12
3
1
0
100
50
0
0°
10°
20°
30°
40°
direction of gaze (eccentricity, deg)

Figure 23.